

UNIVERSIDADE DE LISBOA

Faculdade de Letras

Departamento de Linguística Geral e Românica



**Tradução Automática – proposta para o discurso escrito em língua  
portuguesa**

Natacha Raquel do Carmo Delgado

MESTRADO EM TRADUÇÃO

2013

UNIVERSIDADE DE LISBOA

Faculdade de Letras

Departamento de Linguística Geral e Românica



**Tradução Automática – proposta para o discurso escrito em língua  
portuguesa**

**Natacha Raquel do Carmo Delgado**

Projeto orientado por:  
Prof.<sup>a</sup> Doutora Teresa Seruya

**MESTRADO EM TRADUÇÃO**

2013

## **Agradecimentos**

Gostaria de agradecer, em primeiro lugar, à Professora Doutora Teresa Seruya pelo profissionalismo e interesse incessantes, sem os quais este projeto nunca se teria concretizado. Por me ter incentivado a explorar uma área ainda relativamente desconhecida, cuja aprendizagem considero inestimável. E pela disponibilidade com que orientou este trabalho e, principalmente, por nunca ter desistido de mim.

À minha família, especialmente aos meus avós – pelo amor inextinguível e sacrifício – sem os quais o meu percurso académico não teria sido possível. E ao meu tio, pela generosidade, carinho e apoio ao longo de todo este percurso.

Às minhas melhores amigas, entre os quais destaco a Carla, a Melissa, e a minha prima Ana por acreditarem em mim e nas minhas capacidades. Sem a vossa amizade incondicional teria certamente perdido o Norte, e devo-vos a minha resistência.

Às minhas amigas Teresa e Rita, e aos meus amigos João, André José, Fábio, João Henrique e Marco por toda a paciência, e por toda a força que me deram.

A todas as Professoras e, especialmente, a todos os meninos da escola EB1/JI Brandoa onde leciono – por me terem posto sempre um sorriso nos lábios, com o seu amor, inocência e constante alegria. E por me terem lembrado, diariamente, que a perseverança é recompensada.

Obrigado a todos.

## Resumo

A importância que a tradução automática tem vindo a adquirir, não só pela difusão como também pelo uso crescente de diversas ferramentas e sistemas na área da tradução, tornou necessário que estivesse disponível a maior quantidade possível de informação sobre esta temática. Informação esta que pela importância do seu conteúdo e pela especificidade da linguagem utilizada se torna imperativa estar acessível em língua portuguesa, cuja área começa a ganhar novos contornos.

Como aluna do Mestrado em Tradução considerei que seria um contributo valioso selecionar vários textos importantes na área da TA numa antologia. Dadas as óbvias limitações deste trabalho, uma vez que se trata apenas de um projeto para a obtenção do grau de Mestre, apresento aquilo que seria uma futura antologia de textos cujo conteúdo julgo ser de extrema importância para alunos num contexto semelhante ao meu, facultando mais uma ferramenta de acesso a este tipo de conteúdos. De forma a disponibilizar um número maior de bibliografia em português, proponho tornar disponível, pelo menos, dois dos textos mais importantes da tradução automática, pela relevância histórica, referências e inovação que acrescentam. Esta seleção é acompanhada pela parte do central do presente projeto, i.e. uma análise textual de através do modelo de Christiane Nord, que põe em evidência as características mais importantes de ambos os textos.

Seguem-se em anexo as subseqüentes propostas de tradução, que se pressupõem poder ajudar, também, futuros alunos na compreensão de um conteúdo e de uma linguagem tão complexa.

São linhas orientadoras deste estudo:

1. Seleção de dois textos, a partir de uma lista de, pelo menos, oito obras.
2. Aplicação do modelo de análise textual de Christiane Nord, a partir de *Text Analysis in Translation. Theory, Methodology, and Didactic Application of a Model for Translation-Oriented Text Analysis*.
3. Apresentação das propostas de tradução dos textos escolhidos.

**Palavras-chave:** antologia, Christiane Nord, modelo de análise textual, tradução automática

## Abstract

The increasing importance of machine translation, due not only to the spread and growing use of different tools and systems in translation, made the availability of the largest amount of information on the subject matter necessary. The importance of its content and the specificity of the language used made it imperative for this information to be available in the Portuguese language, giving that the MT field in Portugal has started to gain new outlines.

As a Translation student, I thought it would be a valuable contribution to select several important MT texts in an anthology. Knowing the obvious limitations of this work, and taking into account that this is only a project to obtain my Masters degree, I present what can be seen as a future anthology of texts, the content of which I find extremely important to students in an academic context like my own, providing them an extra tool to access this type of content. In order to provide a bigger number of bibliographies in Portuguese, I propose to make available, at least, two of the most important texts in machine translation for the historic relevance, references and innovation that they present. This selection is accompanied by the main part of the present work, i.e. the textual analysis through Christiane Nord's model, which highlights the most important characteristics of both texts.

As attached documents are the subsequent translation proposals, assuming them to also be helpful to future students who wish to understand such a complex content and language.

The guidelines of this study are:

1. The selection of two texts from a list of, at least, eight publications.
2. Application of Christiane Nord's model of textual analysis, from *Text Analysis in Translation. Theory, Methodology, and Didactic Application of a Model for Translation-Oriented Text Analysis*.
3. The presentation of translation proposals of the chosen texts.

**Keywords:** anthology, Christiane Nord, model of textual analysis, machine translation.

## Índice

<b>1. Introdução.....</b>	<b>7</b>
1.1. Apresentação do tema.....	7
1.2. Breve relatório de investigação.....	8
1.3. Objetivos.....	13
1.4. Metodologia (o modelo de análise textual de Christiane Nord).....	14
1.5. Justificação do <i>corpus</i> .....	17
1.6. Estrutura do trabalho.....	24
<b>2. Aplicação do modelo de análise textual ao texto <i>Translation</i>, de Warren Weaver (1949) .....</b>	<b>25</b>
2.1. Análise dos fatores extratextuais.....	25
2.1.1. <i>Translation</i> .....	26
2.1.2. Reflexo dos fatores extratextuais no texto.....	28
2.2. Análise dos fatores intratextuais.....	31
2.2.1. Análise do efeito.....	38
2.3. Crítica de Tradução.....	39
<b>3. Aplicação do modelo de análise textual ao texto <i>The Development and use of machine translation systems and computer-based translation tools</i>, de John Hutchins (1999).....</b>	<b>50</b>
3.1. Análise dos fatores extratextuais.....	50
3.1.1. <i>The Development and use of machine translation systems and computer-based translation tools</i> .....	51
3.1.2. Reflexo dos fatores extratextuais no texto.....	52
3.2. Análise dos fatores intratextuais.....	53
3.3. Crítica de Tradução.....	65
<b>4. Considerações finais.....</b>	<b>97</b>
<b>5. Referências bibliográficas.....</b>	<b>99</b>
5.1. Bibliografia geral.....	99
5.2. Recursos <i>online</i> .....	102
<b>6. Anexo 1: <i>Translation</i>, de Warren Weaver</b>	
<b>Anexo 2: <i>The Development and use of machine translation systems and computer-based translation tools</i>, de John Hutchins</b>	
<b>Anexo 3: Proposta de tradução para <i>Translation</i>, de Warren Weaver...106</b>	
<b>Anexo 4: Proposta de tradução <i>The Development and use of machine translation systems and computer-based translation tools</i>, de John Hutchins.....</b>	<b>117</b>

# 1. Introdução

## 1. 1. Apresentação do tema

A tradução automática (TA daqui para a frente) sempre foi um recurso fascinante, não só pelo ideal de tradução que apresentava, como também pela inovação que propunha e veio a introduzir na área da tradução. Apesar de todas as incongruências e dos custos dispendiosos que acarreta, a TA veio a adquirir um papel cada vez mais pertinente e indispensável nos dias de hoje.

O progresso tecnológico veio pôr em evidência a potencialidade da TA, permitindo comunicar e partilhar informação em tempo real de forma imediata e eficaz, independentemente da barreira linguística. Apesar de ser uma área que apresenta problemas sistemáticos a nível de funcionalidade<sup>1</sup>, o que é uma óbvia limitação para quem acredita no ideal de tradução automática, o interesse nos recursos que esta área oferece tem crescido cada vez mais. Existe uma investigação cada vez maior e direcionada para várias vertentes da TA<sup>2</sup>, assim como têm tido lugar inúmeros desenvolvimentos que têm ajudado a alargar os seus horizontes e a compreender esta área, até há cerca de duas décadas, largamente desconhecida em Portugal.

Este crescimento da TA em Portugal veio suprir a lacuna de informação que havia sobre o tema. Gradualmente, houve uma necessidade cada vez maior de disponibilizar a maior quantidade de informação possível em português: a história dos sistemas, que sistemas existem, quais os que tiveram mais e menos sucesso, como funcionam, qual o impacto que tiveram na área de tradução, que desenvolvimentos têm sido levados a cabo, entre outros. Como poderemos ver no capítulo “Breve Relatório de investigação”, são vários os profissionais, entre eles professores e investigadores, que se têm dedicado a esta temática. No entanto, grande parte da informação encontra-se disponível em língua inglesa, uma vez que foi em países como os E.U.A. que grande parte da

---

<sup>1</sup> Entre os problemas mais recorrentes e conhecidos está, por exemplo, a incapacidade de um sistema em resolver casos de ambiguidade lexical e estrutural (Somers 2003: 124).

<sup>2</sup> Como é o caso, por exemplo, da tradução de língua falada – *spoken language* – que consiste no processamento computacional da língua falada (havendo a distinção entre língua falada e escrita) (Ranchhod 2001: 195).

investigação que conhecemos hoje teve lugar, e de onde grande parte do poderio tecnológico advém. Contudo, os alunos que, como eu, se interessem pelo tema e que pretendam aprofundá-lo num contexto académico, confrontam-se com uma bibliografia maioritariamente em inglês. Sendo o meio académico um meio privilegiado, onde se podem encontrar inúmeras obras de referência, de que obras o aluno pode dispor sobre os textos mais importantes da TA na sua língua nativa? Que ferramentas de pesquisa pode encontrar e pesquisar em português e que contribuições pode haver nesse sentido?

## **1. 2. Breve relatório de investigação**

Por estarmos perante uma área de conhecimento que reflete o poderio científico e económico do país onde teve lugar, grande parte da bibliografia de TA encontra-se em inglês. No entanto, é importante referir que esta não é desconhecida em Portugal, visto que existem também algumas obras de importante referência em língua portuguesa, embora sejam pouco numerosas em comparação com a produção em língua inglesa. Referir-me-ei neste breve relatório de investigação apenas a publicações posteriores a 2000.

A expansão desta área da tradução nos últimos anos, impulsionada mais recentemente por elementos como a criação e desenvolvimento de dispositivos de tradução *online*, pelo uso cada vez mais frequente de ferramentas de tradução, entre outros, veio influenciar o estado da TA em Portugal e alargar os seus horizontes. Todos estes fatores vieram levantar importantes questões que vários investigadores e professores têm vindo a abordar e a desenvolver nos últimos anos. Ainda que a motivação principal deste trabalho seja contribuir para o enriquecimento da bibliografia em português na área da TA, é importante referir a existência de algumas publicações em língua portuguesa que mostram o desenvolvimento da área em Portugal. Este foi, de resto, o critério de seleção dos textos que irei referir, alguns dos quais foram inclusive consultados para a elaboração do presente projeto.

Entre muito do que existe nos dias de hoje escrito sobre TA em português, é possível encontrar não só livros, como sobretudo teses, artigos publicados para colóquios ou conferências e ainda glossários *online*, que têm vindo a ser desenvolvidos e atualizados permanentemente. Gostaria de começar por referir os livros que se



encontram já publicados, cujo contributo para a área da TA no nosso país foi indispensável. Começo pelo livro de José Manuel Curado, publicado em 2000 e intitulado *O mito da tradução automática*. Este é essencialmente um ensaio que analisa, a nível cultural, os problemas mais recorrentes e mais complexos dos programas computacionais de tradução. Curado foca-se sobretudo nos problemas de contexto e relevância para questionar os problemas inevitáveis destes programas, desmistificando assim o “mito da tradução automática”, que prometia uma tradução automática com uma qualidade igual ou superior à humana. Curado desenvolve os seus argumentos com base em estudos de caso de traduções suas, analisando-os através das teorias de dois investigadores estrangeiros que se encontram em extremidades opostas no debate entre filosofia da mente e filosofia da inteligência artificial. É uma excelente introdução à área da TA, uma vez que fornece várias explicações e põe em evidência indagações pertinentes sobre a TA e o seu papel na área da tradução. Outra das publicações mais importantes de carácter introdutório que podemos encontrar é o *Tratamento de Línguas Por Computador: uma introdução à Linguística Computacional e suas aplicações*, de Elisabete Marques Ranchhod, 2001. Ranchhod, que é de resto uma das grandes especialistas na área da TA e da engenharia da linguagem, edita aquela que é uma obra pioneira no “panorama editorial português” (2001: 9), isto porque a obra se debruça cuidadosamente sobre o tratamento computacional da língua escrita e falada, em particular do português. Sob uma abordagem linguística, Ranchhod edita aquele que é um livro bastante completo e acessível sobre uma temática de difícil compreensão a um leitor não-especialista. O livro aborda questões como a resolução de ambiguidades recorrentes em qualquer língua natural, o uso de dicionários como um recurso linguístico, o PLN, entre outros fatores de extrema importância em TA.

No entanto, uma das obras mais recentes e relevantes é o livro *Avaliação conjunta: um novo paradigma no processamento computacional da língua portuguesa*, de 2007, da autora Diana Santos. Formada em engenharia eletrónica e especialista em TA, Santos tem-se destacado pela investigação e enorme contribuição que tem dado à área. Tem não só disponibilizado vários recursos para a fundação Linateca – conhecido centro de recursos para o processamento computacional da língua portuguesa – como tem elaborado recentemente vários artigos e apresentações em colóquios e seminários, onde tem discutido a importância do computador no que diz respeito à prática da tradução ou às inovações que pretende trazer para o processamento computacional em português. Como o título da sua obra refere, esta reúne profissionais de várias áreas que têm como

objetivo “tornar os computadores fluentes na nossa língua” (2007: 2). Pretendendo que este seja um livro que possa ser lido e compreendido por um leitor que também não possua conhecimentos específicos do domínio, a obra pretende avaliar e facultar pequenas ajudas a qualquer tarefa que inclua a manipulação e tratamento da língua. Entre muitos dos profissionais que contribuíram na elaboração deste livro gostaria de destacar a contribuição da professora Belinda Maia, especialista em linguística computacional e um dos nomes incontornáveis da TA em Portugal. A professora tem inúmeras publicações sobre o tema, muitas das quais se encontram maioritariamente em inglês. No entanto, Maia tem tido colaborações recentes em língua portuguesa, entre as quais se destaca esta obra, para a qual o seu contributo foi, uma vez mais, inestimável.

No entanto, é importante salientar que a TA não trata apenas o processamento de línguas naturais. Engloba questões importantes como as ferramentas de tradução, recursos *online* e a importância cada vez maior da terminologia e da existência de uma base de dados fidedigna e atual. É por este motivo que gostava de chamar a atenção para o livro *Introdução à Terminologia, teoria & prática*, das professoras Maria Krieger e Maria José Finatto. Esta é uma publicação que pretende, precisamente, colmatar a lacuna que existe na produção bibliográfica sobre Terminologia em língua portuguesa (2001: 8). De carácter didático, o livro alerta o leitor para a importância da Terminologia, fornecendo definições básicas e fundamentos teóricos para uma designação que as autoras consideram ser muitas vezes invocada de forma errada. A segunda parte desta importante contribuição para a área da tradução dedica-se tanto às aplicações objetivas, como à redação, à tradução ou à elaboração e gestão informatizada de léxicos especializados.

Nos últimos anos, contudo, os maiores contributos para a área da TA têm sido teses de mestrado e/ou doutoramento, que tentam colmatar a ausência ou insuficiência de determinados campos de conhecimento na área. Gostaria de salientar duas teses orientadas pela Doutora Palmira Marrafa, professora da Faculdade de Letras da Universidade de Lisboa e diretora científica do CLUL – Centro de Computação do Conhecimento Léxico-Gramatical. A primeira, intitulada *Para uma tradução automática baseada em conhecimento: especificação da modificação e predicação adjectiva*, de Náheda Ibrahim, utiliza o lapso que existe entre a adequação linguística e a eficiência dos computadores para examinar o desempenho de ferramentas de TA na *internet* e avaliar a qualidade dos *outputs* em expressões adjetivais em Português-Inglês. Embora a metodologia adotada seja de carácter empírico em sistemas como o Google,

Systran e SDL Free Translation, esta tese vem pôr em evidência muitas das lacunas existentes no estabelecimento de especificações linguísticas nos sistemas, demonstrando como e porque é que ocorrem. Respondendo às próprias questões que lança, Ibrahim afirma que a solução passa inevitavelmente pelo controlo da linguagem, de forma a reduzir a ambiguidade e a complexidade linguística. É nesta medida que a tese de Ana Lucrecia Madeira Gomes, *Tradução Automática e Linguagens Controladas: Contributos para um Português Controlado*, ganha relevância, ao apresentar propostas de uma linguagem controlada – que controla e restringe a língua utilizada – para o português europeu. Esta tese foi um contributo interessante e valioso para a área de TA, até então sem propostas do género. Relativamente recente e importante como contributo foi também a tese de Doutoramento de Ricardo Manuel Brandão Simões, orientada pelo Doutor José João Almeida, *Extracção de Recursos de Tradução com base em Dicionários Probabilísticos de Tradução*, da Faculdade de Engenharia da Universidade no Minho. O objetivo desta tese é encontrar e realizar uma álgebra que permita manusear os recursos bilingues (dicionários de tradução, terminologia bilingue, entre outros), de forma a responder à necessidade de adaptar estes recursos às situações onde se aplicam. Mas não só de teses surgem contributos pertinentes. É frequente nos dias de hoje encontrar artigos que foram apresentados em palestras ou colóquios, que no caso da TA se têm multiplicado pelo país. É caso da XATA, uma conferência nacional anual que tem como tema central *XML: Aplicações e Tecnologias Associadas*, e que tem lugar em várias faculdades. Na conferência de 2007 surgiu o artigo “Processamento de Memórias de Tradução de Grandes Dimensões”, de José João Almeida e Alberto Simões. Este artigo interessa na medida em que analisa a importância das ferramentas CAT (tradução assistida por computador) sob o formato TMX (Translation Memory Exchange) e XML. As ferramentas CAT são uma realidade cada vez mais constante do universo do tradutor. Em Fevereiro de 2007, a investigadora Diana Santos apresentou no Seminário de Tradução Científica e Técnica em Língua Portuguesa, organizado pela União Latina, um artigo intitulado “A tradução na sociedade do conhecimento ou Tradução: uma tecnologia humana de ponta ou Ciência e tradução”. Este é um artigo de opinião, no qual Santos chama a atenção para a importância da tradução como ato construtivo, que age frequentemente como uma ponte entre duas línguas, tendo um objetivo direcional. A tradução, sublinha a autora, é uma tarefa exclusivamente humana que surgiu como resposta a uma necessidade que é compreendida em pleno apenas por humanos. No entanto, o desenvolvimento tecnológico adquiriu um lugar de destaque na

tradução e trouxe com ele várias tecnologias que auxiliam o tradutor, libertando-o de tarefas repetitivas e morosas. Embora não possa substituir o tradutor, a tradução automática tem assumido um lugar de destaque através de inúmeros programas e ferramentas, como a tradução automática baseada em exemplos ou um sistema partilhado de memórias de tradução (Santos 2007: 4). O ponto de maior interesse do seu texto é, contudo, a diferença entre tradução científica e literária. Santos apoia-se em três argumentos importantes para refutar este antagonismo: a demonstração, através de Lackoff e Johnson (1980), de que a língua é essencialmente metafórica, a compreensão de que até algo como a ciência resulta dos cientistas e sua subjetividade, e de como a língua não possui compartimentos para linguagem técnica e geral (Santos 2007: 5). Depois de apresentar os seus argumentos, a autora realça ainda a dicotomia constante entre tradutor e criador, sublinhando a influência de uma língua noutra, e de como é uma tarefa extremamente complexa e árdua saber trabalhar ambas numa equivalência viável.

Além das publicações referidas é essencial complementar qualquer pesquisa tendo presente a importância, não só da terminologia como foi referido, mas também de glossários especializados. É, por conseguinte, necessário lembrar a existência do *Léxico Técnico Português: Ambiente, Banca, Comércio, Construção, Energia, Seguros, Turismo, Telecomunicações, Direito Comercial Internacional e Economia e Gestão de Empresas*, um glossário disponibilizado *online* e elaborado por professores como a Doutora Palmira Marrafa, R. Amaro, R. P. Chaves, S. Lourosa e S. Mendes e do glossário *Visibilidade*, do investigador Daniel Gomes, cujo doutoramento assenta na *web* portuguesa. Gostaria de referir, ainda, o Laboratório de Engenharia da Linguagem, a Languateca e o Centro de Tradumática em Português, do Camões Portugal – Instituto da Cooperação e da Língua, como importantes referências para consultas e investigação na área.

### 1. 3. Objetivos

Por possuir uma forte componente tecnológica, grande parte da bibliografia de TA encontra-se em inglês. É por este motivo que grande parte da terminologia que conhecemos e utilizamos se encontra nesta língua, uma vez que a sua universalidade contribui para a compreensão geral e utilização frequente dos termos criados, refletindo assim o poderio dos países de língua inglesa nesta e outras áreas de dimensão global. Citando Correia “o contexto político e económico mundial, no quadro da globalização, é responsável pela transferência de inúmeros termos técnicos de línguas faladas em estados de maior poderio económico e, conseqüentemente, científico e tecnológico, para línguas de estados menos poderosos.” (Correia 2005: 2). Esta quantidade extraordinária de publicações em inglês veio pôr em evidência o reduzido número de publicações em português, o que no contexto académico como o meu, surge como uma desvantagem. Mesmo existindo um número crescente de publicações em português e um número cada vez maior de investigadores na área, como pudemos observar, um aluno que queira encontrar o máximo de informação possível em português encontra-se condicionado pelo número de textos disponível em língua portuguesa. Esta restrição pode reduzir não só o número de escolhas que um aluno tem, mas também a quantidade de informação que pretende obter de forma clara e objetiva.

O que motivou este trabalho foi, por isso, a verificação enquanto aluna do Mestrado em Tradução (Inglês-Português), de que existe alguma escassez de textos na nossa língua quando em confronto com a língua inglesa. Dado o crescimento gradual da área da TA em Portugal e a importância que tem vindo a assumir progressivamente a saber no ensino da tradução decidi, portanto, minimizar esta lacuna ao apresentar um projeto de uma antologia de textos sobre TA. Pretendo, deste modo, contribuir para o crescimento da bibliografia de TA em português ao disponibilizar uma pequena seleção de textos, e respetivas propostas de tradução, de dois dos nomes mais importantes da sua história. Ao tornar disponível este trabalho pretendo, também, contribuir para o enriquecimento do português como língua científica, uma vez que a seleção é composta por textos científicos e técnicos. Uma tal antologia terá importância académica para alunos que se sintam fascinados pelo tema e, no entanto, têm pouco textos acessíveis em língua portuguesa.

## 1. 4. Metodologia (o modelo de análise textual de Christiane Nord)

Para que a análise apresentada neste projeto fosse rigorosa e solidamente fundamentada foi necessário encontrar uma metodologia de trabalho que se adequasse ao *corpus* e que permitisse pôr em evidência as características mais importantes dos textos escolhidos e, conseqüentemente, as dificuldades e problemas de tradução, bem como as propostas de solução. Um dos pontos mais importantes deste projeto é, por conseguinte, o modelo de análise a propor, que neste caso será o de Christiane Nord.

Em 1988, Nord propõe na sua obra *Textanalyse und Übersetzen: theoretische Grundlagen, Methode und Anwendung einer übersetzungsrelevanten Textanalyse*<sup>3</sup> um modelo de análise de tradução funcional, retomando as posições funcionalistas introduzidas por Vermeer e Reiss<sup>4</sup>. Ao estudar e analisar o ato tradutório e a tarefa do tradutor, Nord verificou que existiam vários elementos e parâmetros que não eram abordados pelos autores referidos ou que se encontravam, segundo a própria, mal categorizados. Era necessário um modelo que permitisse a interpretação correta do texto e explicasse as suas estruturas textuais e linguísticas, de acordo com o sistema e normas do texto de partida (Nord 2005: 1), permitindo assim analisar as diferentes funções, etapas e níveis de importância do processo tradutório para depois os aplicar a qualquer tipo de texto. Nas palavras da autora:

«What is needed is a model of source-text analysis which is applicable to all text types and text specimens, and which can be used in any translation task that may arise. Such a model should enable translators to understand the function of the elements or features observed in the content and structure of the source text. On the basis of this functional concept they can then choose the translation strategies suitable for the intended purpose of the particular translation they are working on. »

(Nord 2007: 1)

---

<sup>3</sup> Para o presente trabalho utilizarei a versão inglesa *Text Analysis in Translation. Theory, Methodology, and Didactic Application of a Model for Translation-Oriented Text Analysis*, de 2005

<sup>4</sup> Noção proposta em: Reiss, K. & Vermeer, H. J. (1984) *Grundlagen einer allgemeinen Translations-theorie*, Tübingen.

É por este motivo que Nord formulou um modelo de análise textual que pudesse não só ser aplicável a todo o tipo de textos e suas particularidades, mas também que fosse específico o suficiente para ajudar a sistematizar o maior número de problemas gerais de tradução. Desta forma, o tradutor conseguiria avaliar de forma consciente todo o processo de tradução e todos os problemas e dificuldades envolvidas, conseguindo justificar de forma clara e ponderada as suas escolhas. É também por esta razão que, como a autora salienta, este é um modelo que tanto pode ser utilizado por um tradutor profissional, como por um estudante ou até por professores cuja função seja preparar e equipar de forma adequada os seus alunos na prática da tradução.

Como foi referido anteriormente, existem para Nord vários parâmetros e fases deste processo que se encontram mal definidos. É por esta razão que a autora sublinha que existe, entre muitos outros exemplos, uma importante diferença entre textualidade e texto, sendo que textualidade diz respeito à estrutura socio-comunicativa dos participantes do ato comunicativo, e texto é a realização concreta da “textualidade” (Nord 2005: 41). Esta distinção feita por Nord reforça um dos critérios principais do seu modelo, no qual o texto é encarado como um ato comunicativo e, como tal, todas as dimensões da situação comunicativa e dos seus participantes são aspetos a ter em conta. Deste modo, todos os fatores que dizem respeito à utilização do texto de partida são de extrema importância para a sua análise, uma vez que determinam a sua função comunicativa (Nord 2005: 41). Nord denomina estes fatores de intratextuais e extratextuais, distinguindo-os para que o tradutor, independentemente do seu nível de experiência, consiga fazer a análise da função do TP e do TC, assinalando todas as dimensões do ato comunicativo sem os confundir ou agrupar sob a mesma categoria.

« From an action-oriented viewpoint, the text is regarded as an element of a communicative interaction which takes place in a situation. The communicative situation (including the participants in the communicative act) becomes then the centre of attention, while the linguistic structure of the text body, which can be analysed by textlinguistic methods, is of secondary importance. It is on the basis of an action-oriented concept of textuality that we may regard the translation of a text as an “action” which makes it possible that a text fulfils certain functions for other participants in a new situation. »

(Nord 2005: 257)

O modelo de Nord apresenta assim duas grandes divisões de análise: os fatores extratextuais (relativos ao emissor, à sua intenção e aos fatores que condicionam a produção do texto) e os fatores intratextuais (relativos ao léxico, estrutura das frases, elementos não-verbais, entre outros). Ambos são essenciais para as fases do processo de tradução, que a autora diz ser um ciclo repetitivo onde o tradutor pode repensar as suas escolhas, e eventualmente modificá-las.

No domínio dos fatores extratextuais importa analisar quem é o autor ou emissor do texto (tendo sempre presente, como sublinha Nord, que existe uma diferença importante entre os dois), qual era a sua intenção quando o elaborou, o meio através do qual foi produzido, onde e quando. Depois de reunidas as respostas a estas perguntas conseguiremos obter aquela que é a resposta à pergunta mais importante da análise dos fatores extratextuais: por que razão foi o texto escrito? (Nord 2005: 42) Todos estes pontos de análise levam-nos a esta pergunta, que Nord diz ser decisiva para a análise da função do texto, uma vez que será a resposta que nos dirá qual o motivo da sua produção. Os fatores intratextuais, por outro lado, são todos os elementos que dizem respeito à constituição intrínseca do texto. Aqui a ênfase é colocada na importância que os elementos linguísticos e verbais assumem, avaliados de acordo com a função que têm no texto. É necessário analisar não só o texto como um todo, mas também os fatores textuais individuais e o seu “ambiente”, uma vez que são estes elementos que se encontram ligados à intenção, função e motivo (fatores extratextuais). O texto é, assim, analisado através da sua microestrutura, ou seja a estrutura formal ou semântica, e através da sua macroestrutura, isto é, as secções em o texto está organizado. Os fatores intratextuais são importantes na medida em que se referem à forma como a informação é apresentada ao recetor, e como é articulada.

A aplicação do modelo inclui três etapas obrigatórias: a análise do perfil do texto de partida, os seus elementos e características que são relevantes e, dentro deste grupo, quais são aquelas que realçam as dificuldades e problemas de tradução que obrigam o tradutor a pensar e repensar as suas escolhas. Nord dá destaque a este percurso para salientar que existe uma importante diferença entre dificuldades (que são subjetivas) e problemas (objetivos) de tradução. As dificuldades de um texto, diz a autora, estão relacionadas com as capacidades do próprio tradutor e as suas condições específicas de trabalho (Nord 2005: 167). Por outro lado, os problemas de tradução estão relacionados com as diferenças específicas de cada língua e cultura, isto porque cada uma denota



hábitos, normas e convenções com respostas e soluções próprias. A distinção entre dificuldades e problemas é importante na medida em que ajudará à sistematização da análise textual, e ajudará o tradutor a justificar as suas escolhas, que resultam do cruzamento entre o vencer das dificuldades e a resolução dos problemas. A justificação das escolhas é, de resto, um dos principais objetivos deste modelo, como já foi referido.

Estamos perante um modelo de análise funcional que pode ser aplicado a qualquer tipo de texto, literário ou não, o que significa que pode ser aplicado com sucesso a um texto científico e/ou técnico, como é o caso do conjunto de textos escolhidos para este projeto. A divisão deste modelo entre os fatores extratextuais e intratextuais é fulcral para analisar o TP e todas as singularidades que o compõem.

« I have attributed great importance to the fact that the model is applicable to source texts of any language and any text type and to translations from and into any language and culture. »

(Nord 2005: 259)

A aplicação do modelo aos textos selecionados terá por base o modelo de Christiane Nord, referido nesta secção, e todos os comentários dos textos-fonte encontrar-se-ão divididos entre fatores extratextuais e intratextuais.

## **1.5. Justificação do *corpus***

O projeto que está por detrás do presente trabalho é a organização de uma antologia em língua portuguesa sobre TA. No entanto, dadas as limitações deste trabalho, analisarei apenas dois textos, de entre os muitos que poderiam figurar num trabalho desta envergadura. Uma antologia ilustra, por definição, uma seleção cuidada de vários textos para determinado público, configurando ou manipulando a sua receção (Seruya 2013: 1).

« Beauty, selection, and re-arrangement are thus qualities and activities present in the genesis of the anthology form. » (Seruya 2013: 2).

Esta seleção é baseada no valor, e envolve um rearranjo e uma reestruturação deliberada de um *corpus*. O objetivo principal de uma antologia é disponibilizar textos, que podem ser marginalizados ou desconhecidos, refletindo e moldando de forma inevitável o gosto literário contemporâneo (Seruya 2013: 3). No entanto, uma antologia pode ter também objetivos educacionais (que têm como função educar os seus leitores ou informá-los sobre determinado tema), de preservação (preservar determinado tipo de bibliografia, sendo a antologia o meio de preservar a memória cultural nacional), inovação (reavaliação de textos e do cânone), proteção (produção de obras literárias menores), estrutura (estruturando e abrindo caminho a um novo ramo na cultura) e acessibilidade (tornando disponível determinada seleção de textos a um vasto público) (Seruya 2013: 3). No caso presente e tendo em conta os objetivos em mãos, a antologia em causa cumpre objetivos educacionais e de acessibilidade.

Como é possível observar, a seleção de textos para uma antologia é uma tarefa que requer muito cuidado, pressupondo a representatividade dos objetos selecionados. Os critérios da sua seleção têm que ser, por isso, rigorosos, embora envolvam alguma subjetividade (Seruya 2013: 4). Contudo, a seleção pode e deve ser determinada pelos seus leitores-alvo, neste caso estudantes de tradução, que podem ter diferentes níveis de conhecimento sobre TA ou tecnologias aplicadas à Tradução.

Por se tratar apenas de um projeto para obtenção do grau de Mestre, o presente trabalho debruçar-se-á apenas sobre dois textos. No entanto, uma futura antologia poderia incluir a seguinte seleção:

- i. ***Translation*, Warren Weaver (1949)**
- ii. ***The Mechanical Determination of Meaning*, Erwin Reifler (1955)**
- iii. ***The Present Status of Automatic Translation of Languages*, Yehoshua Bar-Hillel (1960)**
- iv. ***The Proper Place of Men and Machines in Language Translation*, Martin Kay (1980)**
- v. ***ALPAC: The (In)Famous Report*, John Hutchins (1996)**
- vi. ***The Development and use of machine translation systems and computer-based translation tools*, de John Hutchins**
- vii. ***Translation technologies and minority languages*, Harold Somers (2003)**
- viii. ***Going live on the internet*, Jin Yang and Elke Lange (2003)**

O critério principal de ordenação das publicações acima referidas foi a progressão cronológica entre elas, embora as tenha tentado articular com uma certa progressão temática, desde a história da TA, ao desenvolvimento de sistemas e ferramentas e qual a importância que têm adquirido nos últimos anos. De forma a elucidar o leitor, referir-me-ei brevemente sobre os textos acima mencionados:

**i. *Translation*, Warren Weaver (1949)**

O memorando de Warren Weaver foi o primeiro documento escrito onde a TA foi abordada como uma ideia exequível. Depois de ter trabalhado com o conceito criptográfico, Weaver acreditava que um computador era capaz de efetuar uma tradução através do método criptográfico, sabendo de antemão, pelas limitações que o computador apresentava, que a tradução produzida precisaria sempre de revisão humana. O autor incluiu ainda algumas experiências que fez com colegas seus e possíveis ideias, defendendo sempre que independentemente do resultado obtido, era na TA que se encontrava o futuro da tradução.

**ii. *The Mechanical Determination of Meaning*, Erwin Reifler (1955)**

Publicado poucos anos depois, o texto de Reifler segue o raciocínio do memorando de Warren Weaver. O autor defendia que a TA era a melhor solução tradutória, dado que existiam demasiadas línguas e insuficientes tradutores competentes para conseguir a tradução exequível das mesmas. Pensar em TA, afirmava Reifler, significava contar com as dificuldades e limitações que podiam surgir a nível linguístico e de engenharia. É por este motivo que propôs um sistema de TA que possuísse um rigor semântico e um *output* inteligente o suficiente para que não fosse necessário uma intervenção humana. Para atingir este objetivo, Reifler desenvolveu um sistema que demonstrava a exequibilidade de uma possível mecanização de toda a revisão humana.

**iii. *The Present Status of Automatic Translation of Languages*, Yehoshua Bar-Hillel (1960)**

O artigo de Yehoshua Bar-Hillel é um texto extremamente completo, no qual o autor falou sobre a exequibilidade dos sistemas de tradução automática de qualidade elevada (Fully Automated High Quality Translation – FAHQT), discutindo as razões pelas quais considerava ser um objetivo irreal. A TA, dizia o autor, podia ter objetivos para traduções de baixa qualidade ou parcialmente automáticas. Se fosse esta a premissa, e se houvesse um grande investimento e uma rigorosa divisão de tarefas, seria possível ter, num espaço de 3 a 5 anos, centros de tradução que estivessem aptos a fazer uso da TA a preços competitivos. Bar-Hillel chegou ainda a afirmar que a investigação linguística até então conseguida foi de extrema importância, embora considerasse que não era um requisito fundamental ou particularmente útil, uma vez que não iria trazer a sofisticação pretendida para a TA de qualidade elevada. Grande parte do seu texto retratou ainda a pesquisa feita por vários grupos de investigação, de diferentes países ou dentro do mesmo país (E.U.A.). Grande parte do insucesso verificado resultava, segundo Bar-Hillel, da incompreensão de que era impossível a TA conseguir a sofisticação de uma tradução humana.

**iv. *The Proper Place of Men and Machines in Language Translation*, Martin Kay (1980)**

O texto de Martin Kay fala da importância da tradução e da necessidade de haver o maior número de textos traduzidos. No entanto, disponibilizar traduções competentes em várias línguas e de vários textos é algo que requer tempo e algum investimento. Este é um dos motivos pelos quais a TA foi pensada como uma possível solução, isto porque reduziria os custos que adviriam da contratação de vários tradutores humanos e, certamente, não seria tão morosa. No entanto, Kay salientava que embora tivesse havido uma evolução na ciência da computação, como a análise de algoritmos por exemplo, a TA continuava a ter limitações óbvias a nível linguístico. A solução para este problema, afirmava o autor, seria aproveitar a ideia da TA e os desenvolvimentos adquiridos, para conseguir sistemas de auxílio ao tradutor

v. ***ALPAC: The (In)Famous Report*, John Hutchins (1996)**

Esta é, sem dúvida, uma das publicações mais importantes de Hutchins, uma vez que fala sobre um acontecimento marcante na história da TA: a criação, em 1964, do ALPAC – *Automatic Language Processing Advisory Committee*. Este foi fundamental para a criação e desenvolvimento da linguística computacional e inteligência artificial. No entanto, o relatório ALPAC, conduzido em 1966, concluiu que a TA era lenta, menos rigorosa do que um tradutor humano e, certamente, muito mais dispendiosa. As repercussões foram sentidas, não só a nível económico, uma vez que os Estados Unidos investiram bastante em investigação, mas essencialmente a nível de interesse e resposta. Isto porque os relatórios sobre as incongruências e o insucesso dos sistemas de TA começaram a repetir-se sucessivamente, sem que a raiz do problema fosse identificada. No entanto, como sublinha Hutchins, a existência polémica deste relatório foi essencial para mudar a história da TA, e as abordagens feitas até então.

vi. ***The Development and use of machine translation systems and computer-based translation tools*, de John Hutchins (1999)**

Outro dos importantes contributos para a história da TA foi a elaboração deste artigo de Hutchins sobre os sistemas que foram, desde a sua conceção até à data em que em que o texto foi publicado, mais relevantes e porquê. O autor discute como e em que países determinados sistemas se difundiram. Para além de dar ênfase à proliferação destes sistemas, Hutchins fala ainda da importância que as ferramentas de TA têm adquirido, bem como todas as tecnologias aplicadas à tradução. Esta importância, deve-se, em grande parte, graças ao aparecimento da *internet* e de como permitiu a expansão de vários serviços em rede.

**vii. *Translation technologies and minority languages*, Harold Somers (2003)**

Harold Somers é outro dos nomes importantes na área da TA, e *Computers and Translation: a Translator's Guide* uma das publicações mais recentes e influentes. *Translation technologies and minority languages* é um dos capítulos integrantes da obra referida, e dos poucos da autoria de Somers. O autor fala das ferramentas CAT, i.e. ferramentas de tradução assistida por computador, e de como não existe quase nenhum *software* semelhante para auxiliar o tradutor quando este precisa de trabalhar com línguas menores, como o Hindi ou o Urdu. Somers sugere neste capítulo, por conseguinte, algumas estratégias que ajudariam a minimizar a escassez de ferramentas para estas línguas.

**viii. *Going live on the internet*, Jin Yang and Elke Lange (2003)**

Este é outros dos capítulos que podemos encontrar em *Computers and Translation: a Translator's Guide*. Jin Yang e Elke Lange falam-nos, tal como Hutchins, da importância da *internet*, e de como veio revolucionar a área da TA e da tecnologia. Os dois autores debruçaram-se, especialmente, sobre o aparecimento de vários sistemas de tradução *online*, que prometiam uma tradução rápida e de qualidade elevada. Muitos dos sistemas que conhecemos hoje, como o Babelfish ou o AltaVista, adquiriram a importância que lhes é reconhecida por conseguirem trabalhar os pares de línguas mais procurados. No entanto, os autores fizeram questão de sublinhar que a qualidade que prometiam era, na verdade, bastante inferior, fazendo pouco mais do que uma tradução à letra. Embora conseguisse uma tradução rápida, a qualidade ficava, segundo as empresas que os desenvolveram e os utilizadores, muito aquém do originalmente pretendido.

Desta breve lista foram selecionados para o presente projeto apenas dois textos. E são eles:

▪ ***Translation*, de Warren Weaver.**

De uma importância fundamental na área da TA, o memorando de Warren Weaver foi o primeiro dos textos a ser selecionado para este trabalho e sem o qual esta proposta de antologia estaria incompleta. Conhecido matemático e cientista, Weaver publicou em 1949 aquela que viria a ser uma das publicações mais importantes da TA – senão a mais importante – uma vez que foi a primeira vez que alguém abordou diretamente a possibilidade de uma máquina conseguir fazer uma tradução (até então uma tarefa exclusiva do tradutor humano), com uma qualidade igual à deste. A importância da escolha deste texto recai na sua importância histórica, dado que foi a publicação pioneira desta área onde foram contempladas as primeiras ideias, estudos e experiências sobre o tema. Outros dos motivos relevantes para esta escolha foi a intemporalidade do memorando, uma vez que passados vários anos o ideal de TA continua a ser aquele que Warren Weaver projetou.

▪ ***The Development and use of machine translation systems and computer-based translation tools*, de John Hutchins.**

Importante linguista inglês, John Hutchins é um dos nomes mais relevantes da cena atual da TA. A contribuição de Hutchins, tanto a nível de investigação, como a nível de obras escritas sobre o tema destaca-se e, como tal, era imperativa a inclusão de, pelo menos, uma das suas publicações no presente projeto.

Especialista no uso de computadores em tradução, Hutchins começou por interessar-se pelos sistemas que existiam, compilando-os em vários livros, sem deixar de dissertar sobre as funcionalidades de cada um e qual a sua difusão e *feedback*. As publicações do autor pautam-se sempre por um breve percurso histórico, de forma a tornar o leitor (erudito ou não) consciente da evolução dos sistemas mais relevantes da área. Achei, portanto, pertinente a inclusão desta publicação no *corpus* de trabalho, uma vez que completa a informação que o memorando de Weaver introduz, apresentando

mais informação à história dos sistemas e referindo quais os desenvolvimentos e progressos que foram feitos em TA nos últimos anos.

## **1. 6. Estrutura do trabalho**

O presente projeto apresenta, além da Introdução, Conclusão e Referências Bibliográficas, a aplicação do modelo de análise de Christiane Nord aos textos acima mencionados, de acordo com a metodologia proposta. É importante referir que tanto as publicações selecionadas, como as respetivas propostas de tradução se encontram em anexo (anexo 1 e 2 respetivamente) na parte final deste trabalho.



## 2. Aplicação do modelo de análise textual ao texto *Translation*, de Warren Weaver (1949)

### 2.1. Análise dos fatores extratextuais

Conhecido matemático e cientista, Warren Weaver (1894-1978) (AUTOR DO TEXTO e EMISSOR<sup>5</sup>)<sup>6</sup> esteve, desde muito cedo, ligado às ciências aplicadas e à matemática. Embora planeasse uma carreira ligada à área da engenharia, a influência dos professores Max Mason e Charles Slichter (ambos com uma longa e prestigiada carreira como matemáticos) levou Weaver a interessar-se pela matemática ao ponto de mudar de licenciatura<sup>7</sup>. Depois de cumprir serviço militar na Força Aérea (1917-19), Weaver deu aulas como professor assistente de matemática no Instituto Politécnico de Throop, entre 1917 e 1918, antes de regressar à Universidade de Wisconsin e de se ter doutorado em 1921 na mesma área. Não só chegou a dar aulas como professor na mesma Universidade, como veio, inclusive, a tornar-se presidente do departamento de matemática entre 1928 e 1932. O profissionalismo e a boa relação que manteve com Max Mason, levou a que fosse convidado a ser diretor do departamento de Ciências Naturais na Fundação Rockefeller, depois de Mason se ter tornado presidente em 1929. Weaver foi convidado para se tornar responsável e incentivar vários projetos que tinham como finalidade o progresso tecnológico através do intercâmbio com outros campos de estudo, função que desempenhou com distinção e na qual se destacou<sup>8</sup>.

---

<sup>5</sup> Embora este texto tenha sido reimpresso com a autorização do autor e publicado mais tarde em Locke, William N. & Booth, A. Donald (1955) *Machine translation of languages: fourteen essays*, Technology Press of the Massachusetts Institute of Technology, Cambridge, Mass., and John Wiley & Sons, Inc., New York, o memorando de Weaver foi originalmente enviado pelo próprio autor a colegas seus. A verificação deste facto encontra-se na nota de editor, presente no final do texto. Este pode ser encontrado no anexo 1.

<sup>6</sup> A utilização de parêntesis com maiúsculas respeita as convenções do modelo de análise textual de Nord, pelo que será aplicada, daqui para a frente, a ambos os textos seleccionados.

<sup>7</sup> Factos relativos à biografia do autor podem ser encontrados em: <http://www-history.mcs.st-and.ac.uk/Biographies/Weaver.html>

<sup>8</sup> Segundo os relatos de um colega. Em Reingold, N. & Warren Weaver, *American National Biography*, 22 (Oxford, 1999), 838-841. Acedido em <http://www-history.mcs.st-and.ac.uk/Biographies/Weaver.html>

Contudo, foi na Segunda Grande Guerra, ao ser destacado para o painel de Matemáticas Aplicadas do governo americano, que Warren Weaver se familiarizou com máquinas eletrônicas e o seu desenvolvimento. Isto permitiu que pudesse estudar e, principalmente, aplicar várias técnicas matemáticas e de estatística a estas máquinas para conseguir decifrar mensagens inimigas (MOTIVO). É exatamente esta a premissa do seu memorando, no qual Weaver, ao perceber que vários métodos científicos podiam ser aplicados a dispositivos eletrônicos e a decodificar mensagens simples como acabara de acontecer durante a Segunda Grande Guerra, se alicerça no conceito matemático da criptografia para fazer a analogia entre a tradução e a decodificação (INTENÇÃO).

## **2. 1.1. *Translation***

O texto de Warren Weaver, intitulado simplesmente de *Translation* [Tradução], é um memorando que o próprio autor elaborou em Julho de 1949 (TEMPO da produção do texto), e enviou a cerca de 200 colegas seus, cientistas e matemáticos (PÚBLICO). Esta é uma publicação breve que aborda essencialmente quatro propostas de temas inovadores dentro do campo da tradução (INTENÇÃO). Warren Weaver começa por apresentar-nos aquele que acredita ser o conceito-chave da sua problematização: a criptografia. A introdução deste tema surge através da experiência pessoal de um colega matemático com a língua turca e a dificuldade que este tinha em compreendê-la. Tal como Weaver, este seu colega partilhava o mesmo interesse pela criptografia (MOTIVAÇÃO) e começou a decifrar mensagens simples, de forma a poder compreender finalmente a língua e, por conseguinte, a exequibilidade de uma possível decodificação. De forma a dar pertinência a este conceito e ao seu uso, Weaver utiliza a sua própria experiência com o conceito criptográfico (CONTEÚDO) durante a Segunda Grande Guerra para justificar a sua utilização dentro da área da tradução. Esta introdução através de encontros pessoais, conversas ou troca de correspondência é, de resto, uma das particularidades mais marcantes do texto, e da própria escrita de Weaver, sempre que fala de um assunto inovador. (COMPOSIÇÃO).

Outro dos temas que o autor aborda prende-se com a discussão da capacidade dos computadores em conseguir uma mecanização do discurso bem-sucedida, lidando e resolvendo simultaneamente problemas de polissemia, semântica, entre outros, até então resolvidos com sucesso apenas por tradutores humanos. Este tema é mais uma vez introduzido através de uma experiência do próprio, nomeadamente uma troca de correspondência entre Weaver e o professor Robert Wiener, famoso matemático e preconizador da cibernética. Depois de alguma reflexão Weaver decide enviar-lhe uma carta na tentativa de interessar o professor, tão devoto a desenvolver as teorias de robótica, automatização, entre outras, a resolver o problema da tradução automática. Contudo, Wiener recusa-se a acreditar que uma máquina como um computador conseguisse solucionar os problemas de barreiras linguísticas que advêm das singularidades de cada língua. É aqui que Warren Weaver descobre que o seu ideal de tradução automática já tinha sido abordado antes pelo professor Andrew Booth, da Universidade de Londres. Decidido a partilhar e aprofundar ainda mais este tema, Weaver decide encontrar-se com Booth e ainda Richard H. Richens, director-assistente da Agência de Multiplicação de Plantas e Genética e também ele fascinado pelo problema de tradução, para juntos testarem a exequibilidade dos seus ideais automáticos. A inclusão deste encontro no seu memorando é de extrema importância, dado que é através desta pequena reunião que Weaver consegue partilhar todo o conhecimento que possui (e que adquirira também durante a Segunda Grande Guerra) e que, juntamente com Richens e Booth, explora a nível prático. Os três propõem uma tradução “rascunho” à letra, na qual a máquina procurava encontrar determinada palavra e caso esta não constasse da memória da máquina, descartava a última letra e tentava de novo.

Um dos grandes erros desta proposta conjunta é ter-se limitado à funcionalidade da máquina, sem considerar os possíveis problemas que normalmente advêm de qualquer acto tradutório: problemas como o da ambiguidade, por exemplo. Como Weaver sublinha, a preocupação na altura fora única e exclusivamente o processo de mecanização do dicionário, que obviamente não fora suficiente para colmatar as falhas que advieram. Contudo, os autores puderam concluir a partir desta experiência aquilo que todos estes anos de história e progresso dos sistemas de tradução automática nos permitem facilmente concluir nos dias de hoje: que uma tradução automática jamais pode trabalhar com traduções literárias, uma vez que os problemas que podem advir são

mais frequentes, complexos e passam por soluções que só um tradutor humano tem capacidade para resolver. As traduções técnicas, por sua vez, já possuem problemas de ambiguidade mais simples de resolver, podendo dar concretização aos ideais de descodificação que propõem.

É por esta razão que Warren Weaver salienta os seus dois últimos tópicos. Sabendo que o conceito de criptografia é de cariz matemático, Weaver apoia-se na “teoria da informação” (CONTEÚDO, MOTIVAÇÃO) – conceito desenvolvido por Claude Shannon – para reforçar um dos fatores mais importantes a ter em conta: a linguagem tem um carácter estatístico e os problemas que dela advêm podem ser resolvidos como se de um teorema ou probabilidade se tratasse. Debruçando-se um pouco mais sobre o contexto e significado de determinado conteúdo, Weaver relembra que a linguagem dentro do sistema tem propriedades estatísticas e quaisquer que sejam os problemas que a tradução encontre, pode resolvê-los da mesma forma que conseguiu efectuar a tradução: através de fórmulas e propriedades estatísticas (INTENÇÃO, MOTIVAÇÃO). O autor dá o exemplo do problema da ambiguidade, lembrando que quanto maior for o leque de traduções possíveis, maior a lentidão da máquina em resolver o problema. Esta probabilidade pode aumentar ou ser reduzida consoante a análise do computador.

É por a linguagem conter elementos “ilógicos” que a criptografia surge como a melhor ferramenta da tradução, uma vez que não garantindo uma tradução perfeita garante, pelo menos, uma tradução com uma margem de erro limitada. Por último, Weaver faz a ponte entre a utilidade deste conceito e os vindouros (e utópicos) resultados. Embora a tradição mítica reporte a origem da tradução à Torre de Babel, Weaver propõe através da sua abordagem criptográfica uma espécie de torre “anti-Babel”. Esta abordagem facilitaria, segundo ele, a comunicação, facultando estruturas comuns e permitindo aos seus intervenientes escolherem o caminho que lhes aprouvesse (INTENÇÃO) – sem que para isso tivessem de “gritar” de cada uma das suas “torres”.

### **2.1.2. Reflexo dos fatores extratextuais no texto**

Uma das particularidades deste texto é o facto de o autor (AUTOR DO TEXTO e EMISSOR) se referir a si próprio na terceira pessoa, algo comum em textos mais científicos, como que uma estratégia de redução de subjetividade.

- “(...) it was very natural **for W. W. to think**, several years ago, **of the** possibility (...)” (Weaver 1949: 14).

Embora existam casos pontuais em que o autor se exprime na primeira pessoa.

- “**This may be, for all I know, a famous theorem of philology**” (Weaver 1949: 13).

Por ser um texto de cariz mais científico, o texto é marcado por uma linguagem maioritariamente formal, uma vez que existem expressões de distinção social (“Professor”, “Dr.”, entre outros) e de elucidação face às funções e especialidades das pessoas descritas. Exemplos como o indicado abaixo ditam o tom do texto e são, também, normalmente típicas deste tipo de texto:

- “On May 25, 1948, **W. W. visited Dr. Booth in his computer laboratory at Welwyn, London, and learned that Dr. Richens, Assistant Director of the Bureau of Plant Breeding and Genetics** (...)” (Weaver 1949: 15)

No entanto, a linguagem utilizada nem sempre é formal e existem alguns casos em que o autor se aproxima do leitor em pequenas observações ou comentários, em tom coloquial e próprio da oralidade:

- “The suggestions of this memorandum **will surely be incomplete and naïve**, and **may well be patently silly to** an expert in the field (...)” (Weaver 1949: 13).
- “(...) **one naturally wonders** if the problem of translation could conceivably be treated as a problem in cryptography.” (Weaver 1949: 14)

- “First, **let’s us think of a way** in which the problem of multiple meaning can, **in principle at least**, be solved.” (Weaver 1949: 15)

Sendo que o estilo de um texto é a forma pela qual o pensamento é expresso, o memorando de Weaver é caracterizado por elaborações intelectuais e científicas – das propostas apresentadas e desenvolvidas (FUNÇÃO DO TEXTO). Dado que a sua publicação tem como finalidade a divulgação de conceitos e propostas, nas quais Weaver acredita residir o futuro da tradução, o texto é maioritariamente descritivo/expositivo. Esta função informativa existe para que o leitor possa perceber não só qual a intenção, mas em que é que se baseia e que fundamentos têm as suas ideias (relação entre EMISSOR e RECETOR). Faço aqui um pequeno aparte para referir que ambos os textos selecionados têm muitas características em comum. No entanto, é importante salientar que um dos textos é de cariz mais científico que o outro, nomeadamente o texto agora em análise.

Consideram-se próprias de um texto mais científico as seguintes características:

- a) Termos técnicos: *mathematical theory of communication, mechanization of speech, statistical semantic*, entre outros (LÉXICO)
- b) Uso do presente como tempo verbal nos comentários sobre os desenvolvimentos, ideias e encontros que teve:
  - “There **is** no need to do more than **mention the obvious** fact that a multiplicity of languages **impedes** (...)” (Weaver 1949: 13)
  - “(...) **there is** a great deal of evidence **for this** general viewpoint.” (Weaver 1949: 14) (ESTRUTURA DAS FRASES)
- c) Dimensão retórica e argumentativa:
  - “Also knowing nothing official about, but having guessed and inferred considerable about, powerful new mechanized methods in cryptography—methods which I believe succeed even when one does not know what language has been coded (...)” (Weaver 1949: 13) (COMPOSIÇÃO, FUNÇÃO DO TEXTO)
- d) Nomes de pessoas e/ou entidades e fontes: *UNESCO, Bureau of Plant Breeding and Genetics, Michael Shannon, Professor Wiener, Bell Telephone Laboratories*, entre outros (CONTEÚDO)

e) Citações diretas e indiretas como:

- “Reichenbach remarked, ‘‘I was amazed to discover that, for (apparently) widely varying languages, the basic logical structures have important common features.’’ (Weaver 1949: 14) (CONTEÚDO e COMPOSIÇÃO)

A apresentação do conteúdo é devidamente estruturada em subdivisões que estão interligadas de forma lógica demonstrando, por conseguinte, a articulação dos temas propostos com o conceito de tradução (COMPOSIÇÃO e CONTEÚDO). Esta estrutura permite a progressão textual do raciocínio que Weaver apresenta, e que pretende ser considerado uma opção científica viável.

## **2. 2. Análise dos fatores intratextuais**

### **a) Assunto**

De acordo com as convenções do que constitui um texto científico e técnico, o título do texto e os subtítulos que possui indicam, frequentemente, qual o(s) tema(s) principal(ais) do texto. Porém, o memorando de Warren Weaver possui, ao contrário do que se espera deste tipo de texto, um título bastante genérico. Intitulado simplesmente de *Translation*, o cabeçalho do texto de Weaver está longe de demonstrar as ideias e propostas que as suas subdivisões apresentam. É a minha opinião pessoal que a escolha de um título tão abrangente como *Translation* foi a tentativa de Weaver em conseguir alguma atenção por parte dos seus leitores. Sabendo que propunha ideias pioneiras e vanguardistas na área da tradução, e sabendo que haveria à partida quem as rejeitasse (como é o caso do Professor Wiener, cuja opinião é inclusive incluída no memorando), penso que Weaver terá decidido apostar num título que não fosse polémico e que não levasse à possível rejeição da sua leitura. Ao utilizar um título tão genérico como *Translation*, Weaver garante que ninguém vai declinar *a priori* as suas ideias, e consegue suscitar uma certa curiosidade no leitor, levando-o a querer saber que assunto(s) relacionado(s) com tradução vai abordar ao certo. A

expectativa que o título suscita é depois confirmada pelos elementos isotópicos do texto. São apenas as subdivisões da publicação de Weaver que estão de acordo com as características do que constitui um texto científico e técnico e indicam especificamente que tema vai ser abordado.

## **b) Conteúdo**

A análise do conteúdo demonstra, ao mesmo tempo, a macroestrutura do texto:

**1. Pequena introdução:** Onde o autor apresenta o tema e justifica a pertinência da elaboração do memorando, e da importância das sugestões que nele figuram.

**2. A War Anedocte – Language Invariants:** introdução do conceito da criptografia através da experiência de um colega seu com a língua turca, e de como foi possível decodificar mensagens de turco para inglês, sendo esta a premissa para a viabilidade do conceito da criptografia.

**3. Translation and Computers:** discussão sobre a capacidade de um computador efetuar uma tradução de qualidade aceitável, através da troca de correspondência entre Weaver e o Professor Wiener, que o autor acreditava ser a pessoa indicada para explorar e desenvolver este conceito. Não encontrando apoio em Wiener, Weaver encontra-se com o professor Booth e com o professor Richens para elaborar uma proposta de tradução, que suscita importantes lacunas, nomeadamente a nível da ambiguidade. Isto permitiu concluir que a delimitação ao tipo de tradução ajuda a diminuir o nível de incoerências encontradas.

**4. The Future of Computer Translation:** refere que não existe qualquer indicação promissora de quais podem ser, no futuro, as possibilidades para a TA. O autor refere ainda que cabe a alguém com as devidas competências assegurar quais as hipóteses e a relevância das mesmas.

**5. Meaning and Context:** O autor debruça-se sobre o problema mais grave e recorrente da TA, a questão da ambiguidade e de como pode, segundo ele, ser resolvida. A ambiguidade prende-se, segundo Weaver, com o carácter semântico e estatístico da linguagem. Se for escolhido um artigo técnico ou científico a probabilidade de o computador encontrar um termo ambíguo diminui, sendo que



a solução passa por tentar minimizar a probabilidade estatística de termos ou expressões dúbias.

**6. Language and Logic:** referência ao teorema de Pitts e McCulloch, de 1943, onde foi constatado que um computador, se concebido com determinados *loops* regenerativos, ao fim de algum tempo é capaz de inferir conclusões legítimas. Isto é, se não estiver a lidar com uma tradução literária pois, neste caso, os problemas em mãos são maiores e as conclusões que podem ser retiradas serão subjetivas.

**7. Translation and Cryptography:** de seguida, o autor chama a atenção para a teoria matemática da comunicação, formulada por Claude Shannon. Esta teoria é muito importante para o conceito da criptografia, uma vez que trata o problema da TA do ponto de vista semântico e estatístico.

**8. Language and Invariants:** Na última subdivisão, Weaver apresenta a TA como uma forma de derrubar a secular Torre de Babel, sublinhando que, mesmo que não se chegue a resolver o problema da automatização da tradução, o processo comunicativo sairá facilitado.

**9. Editor's Note:** Uma nota de editor, que contém a data e o ano em que o texto foi escrito, e onde o editor sublinha que o memorando foi impresso com a permissão do autor, uma vez que é um documento histórico sobre a TA.

**10. References:** à teoria de Pitts e McCulloch, e à versão simplificada da Teoria Matemática da Comunicação

### **c) Pressuposição**

É importante sublinhar que uma das características mais relevantes do texto é o facto de o autor pressupor que os seus 200 colegas possuem os conhecimentos indicados para compreender e aceitar o que propõe. Weaver conta com isso para argumentar a favor das suas ideias, apesar de estar ciente das desvantagens que possam existir. É também por este motivo que é importante mencionar que este memorando foi enviado pelo próprio autor a colegas seus, cientistas e matemáticos – o que significa que o autor avaliou muito bem a situação do seu texto, teve em conta o conhecimento do recetor e soube aprofundar a relevância da informação que pretendia transmitir.

Existem, por conseguinte, vários elementos lexicais que o autor utiliza que demonstram isso mesmo. Entre eles:

- i. “And, in any event, **there are obvious reasons** which make the supposition a likely one.” (Weaver 1949: 13)
- ii. “**It is, in fact, amply clear** that a translation procedure that does little more than handle a one-to-one correspondence of words **cannot hope to be useful** for problems of literary translation (...)”. (Weaver 1949: 15)
- iii. “First, **let us think of a way** in which the problem of multiple meaning can, in principle at least, be solved. **If one examines the words in a book** (...)”. (Weaver 1949: 15)
- iv. “**Thus one is led to** the concept of a translation process in which, in determining meaning for a word, account is taken of the immediate (2N-word) context”. (Weaver 1949: 16)

#### **d) Composição**

Como já referido, o texto em análise apresenta-se estruturado em sete subdivisões que exprimem a progressão lógica das ideias que Warren Weaver apresenta e pretende desenvolver.

Cada subdivisão apresenta uma nova ideia ou o desenvolvimento desta, havendo sempre uma correlação entre cada subdivisão.

Dado que um texto de cariz científico possui uma dimensão retórica e argumentativa, a sua composição apresenta, regra geral, elementos que sustentem uma teoria (Fontanet 2011: 2). É por este motivo que à medida que o autor introduz um novo conceito que pretende desenvolver, ou que quer que o leitor entenda porque é que é relevante, sustenta-o através de exemplos como experiências, trocas de correspondência com colegas seus e referências a teorias que suportam as suas convicções. À exceção de duas subdivisões, todas as restantes argumentam na base de exemplos concretos.

A composição do memorando é, deste modo, pautada pela articulação entre a exposição/articulação da temática, as referências que sustentam a validade do que propõe e os elementos do texto que revelam o seu carácter argumentativo. É possível verificar isso em exemplos como:

- i. “Having had considerable exposure to computer design problems during the war, and being aware of the speed, capacity, and logical flexibility possible in modern electronic computers, it was very natural for W. W. to think, several years ago, of the possibility that such computers be used for translation. On March 4, 1947, after having turned this idea over for a couple of years, W. W. wrote to Professor Norbert Wiener of Massachusetts Institute of Technology as follows:” (Weaver 1949: 14)
- ii. “It is, in fact, amply clear, that a translation procedure that does little more than handle a one-to-one correspondence of words cannot hope to be useful for problems of *literary* translation, in which style is important, and in which the problems of idiom, multiple meaning, etc., are frequent.” (Weaver 1949: 15)

**e) Elementos não-verbais**

O primeiro elemento não-verbal que salta à vista e que é, de resto, característico da maioria dos textos científicos e técnicos é o espaço que existe entre cada subdivisão do texto (Nord 2005: 118). A razão para este *layout* prende-se com a sua composição, e reside na evidente mudança de tema, ou na interligação do tema anteriormente descrito com o que pretende abordar.

Nos exemplos que o autor utiliza para ilustrar os seus argumentos, salta também à vista o tamanho da letra e o espaço utilizado na transcrição do excerto relativo à troca de correspondência com o professor Wiener. Esta diferença de letra obriga o leitor a focar a sua atenção nesta troca de correspondência pessoal, prendendo a sua atenção e levando-o, mais tarde, a recordar-se dela para fazer uma associação lógica de ideias.

**f) Léxico**

O léxico utilizado num texto, sendo influenciado por fatores intra – e extratextuais, é revelador da intencionalidade desse texto (Nord 2005: 122). Por ser um texto pioneiro na área da TA, não existe uma grande quantidade de léxico terminológico. E dado o tempo de produção do texto – 1949 – o léxico sobre TA que podemos encontrar é menor e menos complexo que o que existe nos dias de hoje.

Entre o léxico encontrado destacam-se termos como:

- *Decoding*
- *Computer design*
- *Cybernetics*
- *Mechanization of speech*
- *Quasimechanical*
- *Statistical semantic*
- *Mathematical Theory of Communication*
- *Regenerative loops*

**g) Estrutura das frases**

Dada a apresentação dos elementos não-verbais e a estrutura adotada neste texto, as construções sintáticas são geralmente mais longas do que o habitual em língua portuguesa. O autor começa por apresentar o conceito de forma sucinta, através de frases breves e só depois faz uso de construções mais longas para elaborar sobre o que pretende.

Pode também dizer-se que esta estrutura longa das frases é própria do raciocínio científico, permitindo acrescentar várias informações sobre o conteúdo pretendido e conferindo-lhe a devida complexidade. São exemplos:

- i. “The idea of basic common elements in all languages later received support from a remark which the mathematician and logician Reichenbach made to W. W. Reichenbach also spent some time in Istanbul, and, like many of the German scholars who went there, he was perplexed and irritated by the Turkish language.” (Weaver 1949: 14)

- ii. “On May 25, 1948, W. W. visited Dr. Booth in his computer laboratory at Welwyn, London, and learned that Dr. Richens, Assistant Director of the Bureau of Plant Breeding and Genetics, and much concerned with the abstracting problem, had been interested with Dr. Booth in the translation problem.” (Weaver 1949: 15)
- iii. “The formal truth of this statement becomes clear when one mentions that the middle word of a whole article or a whole book is unambiguous if one has read the whole article or book, providing of course that the article or book is sufficiently well written to communicate at all.” (Weaver 1949: 16)

#### **h) Características suprasegmentais**

O texto contém ainda vários elementos formais da linguagem, que ajudam a organizar verbalmente o que o autor pretende exprimir. Estes elementos expressam o “tom” que o autor adotou e a entoação com que se dirige ao leitor, como se de um diálogo se tratasse.

Outro dos elementos relativos às características suprasegmentais do texto é a ênfase dada a determinadas palavras ou frases, que o leitor reconhece inconscientemente através do uso recorrente de itálico por parte do autor. Exemplos disso são alguns dos comentários e referências que faz, ou a nota de editor no final do texto com uma letra de tamanho mais pequeno.

Destaquei a negrito algumas expressões que denotam qual a entoação e quais as marcas de conhecimento e formalidade no texto. São características suprasegmentais elementos como:

- i. “The present memorandum, **assuming the validity and importance** of this fact, contains some comments and suggestions **bearing on the possibility of contributing** at least something to (...)”. (Weaver 1949: 13) - formalidade
- ii. “During the war **a distinguished mathematician** whom we will call P (...)” (Weaver 1949: 13) – formalidade
- iii. “**Also knowing nothing official about, but having guessed and inferred considerable about**, powerful new mechanized methods in cryptography (...)”. (Weaver 1949: 14) – formalidade

- iv. “(...) **but one could hardly avoid guessing** that this process made use of frequencies of letters, letter combinations (...)”. (Weaver 1949: 13) – marca conhecimento do leitor
- v. “This may be, **for all I know**, a famous theorem of philology.” (Weaver 1949: 13) – entoação
- vi. “Thus their interest was, at least at that time, confined to the problem of the mechanization of a dictionary which in a reasonably efficient way would handle *all forms* of all words.” (Weaver 1949: 15) – entoação
- vii. “Or would it determine unique meaning in a satisfactory fraction of cases, to examine not the  $2N$  adjacent *words*, but perhaps the  $2N$  adjacent *nouns*? What choice of adjacent words maximizes the probability of correct choice of meaning, and at the same time leads to a small value of  $N$ ?” (Weaver 1949: 16) – entoação

### 2.2.1. Análise do efeito

O efeito é uma categoria de análise bastante importante, uma vez que se refere à relação do texto com os seus utilizadores. Esta análise tem que ser observada do ponto de vista do recetor, uma vez que é depois de receber o texto que este vai comparar a sua expectativa com o seu conhecimento. O efeito de um texto é, tal como os restantes elementos relativos ao texto, determinado pelos fatores extratextuais e intratextuais. Existem três tipos de efeito: imediato, curto e a longo prazo.

Os elementos do texto que nos permitem identificá-lo como um texto de cariz ligeiramente mais científico – a dimensão argumentativa, o léxico, a inclusão de inúmeras referências – asseguram a veracidade e a viabilidade da informação (EFEITO INTENCIONAL). Estando ciente de que está a propor ideias vanguardistas que o leitor, erudito, pode e irá inicialmente refutar, o efeito do texto de Weaver é a longo prazo. Assume-se, pela inclusão de pelo menos um exemplo em que um colega seu descartou a sua ideia (Weaver 1949: 14-15), que não iria conseguir um efeito imediato nos seus leitores. Dado que estamos perante um texto de 1949, que trata assuntos que implicam alguma investigação e investimento, o efeito também não terá sido a curto prazo.

No entanto, sabe-se que o professor A. D. Booth, um dos nomes referidos por Weaver e que chegou inclusive a colaborar consigo em algumas ideias relacionadas com TA, escreveu alguns anos mais tarde um artigo intitulado *Mechanical Translation*<sup>9</sup>, em 1955. Neste texto, o professor dá continuação às propostas de Weaver, aprofundando a ideia da tradução automática e de como poderia ser exequível, demonstrando que a publicação de Weaver teve, pelo menos em Booth, um efeito a longo prazo.

## 2.3. Crítica de Tradução

### a) Coerência

A coerência de um texto é importante, na medida em que deixa claro para o seu leitor a relação entre os seus elementos. A compreensão da informação que é verbalizada no texto é conseguida através da coordenação do leitor, coordenação esta que é ativada pela sua memória armazenada (competência linguística, situação comunicativa, proficiência do uso da língua, “horizonte” do recetor) (Nord 2005: 143).

No texto de Weaver esta coerência é conseguida através da inclusão de correspondência, reuniões e referências bibliográficas que vêm sempre datadas. Isto confere uma progressão cronológica, permitindo ao leitor situar-se no tempo e seguir o raciocínio que o autor vai desenvolvendo em cada subcapítulo.

Exemplos:

- i. **“During the war, when the whole field of cryptography was so secret, it did not seem discreet to inquire** concerning details of this story; but one could hardly avoid guessing that this process made use of frequencies of letters, letter combinations, intervals between letters and letter combinations, letter patterns, etc., *which are to some significant degree independent of the language used.*” (Weaver 1949: 13) // **“Durante a guerra, quando toda a área da criptografia era tão secreta, parecia indiscreto inquirir** os detalhes relativos a esta história; mas não é difícil adivinhar que este

---

<sup>9</sup> O texto foi publicado em *Computers and Automation*, vol., 2, no. 4, 1955. Numa bibliografia mais acessível pode ser encontrado em: Nirenburg, Sergei (2003) *Readings in Machine Translation*, Cambridge, Massachusetts, London, England, Bradford Book

processo fez uso da frequência de letras, da combinação de letras, intervalos entre letras e combinações de letras, padrões de letras, entre outros, *que são em larga medida um grau independente da língua utilizada*.

- ii. **“On March 4, 1947, after having turned this idea over for a couple of years, W. W. wrote to Professor Norbert Wiener of Massachusetts Institute of Technology as follows:”** (Weaver 1949: 14) // **“A 4 de Março de 1947, depois de ter dado voltas a esta ideia durante alguns anos, W.W escreveu ao professor Norbert Wiener do MIT (Massachusetts Institute of Technology) nos seguintes termos:”**
- iii. **“The idea has, however, been seriously considered elsewhere. The first instance known to W. W., subsequent to his own notion about it, was described in a memorandum dated February 12, 1948, written by Dr. Andrew D. Booth, who, in Professor J. D. Bernal’s department in Birkbeck College, University of London (...).”** (Weaver 1949: 15) // **“A ideia já foi, contudo, considerada anteriormente. O primeiro exemplo do conhecimento de W.W subsequente à sua própria noção do assunto, foi descrito num memorando datado de 12 de Fevereiro, de 1948, escrito pelo Dr. Andrew D. Booth que, no departamento do Professor J. D. Bernal, na Birkbeck College University of London [Faculdade de Birbeck da Universidade de Londres] (...).”**
- iv. **“A more general basis for hoping that a computer could be designed which would cope with a useful part of the problem of translation is to be found in a theorem which was proved in 1943 by McCulloch and Pitts.”** (Weaver 1949: 16) // **“Uma base mais geral para esperar que um computador possa ser concebido, de forma a lidar com uma parte útil do problema da tradução, é encontrada num teorema, testado em 1943 por McCulloch e Pitts.”**

Embora o autor contribua com várias referências, às quais foi difícil encontrar um equivalente em português, como por exemplo “Bureau of Plant Breeding and Genetics” que não tem uma tradução estabelecida, a inclusão destes elementos não se tornou problemática aquando da tradução. A inclusão de datas e nomes facilitou a progressão cronológica da proposta de tradução e permitiu que pudesse repensar algumas das



minhas opções tradutórias, de forma a conferir o mesmo “tom” do autor e tendo em conta o público-alvo original.

### **b) Estilo Científico**

Como já referido, uma das especificidades de um texto que possui algumas características deste cariz é o facto de o autor se incluir nos exemplos que ilustra e se referir a si próprio na terceira pessoa. Em português, o mesmo pode ser feito sem levantar problemas. No entanto, existem muitas construções sintáticas em inglês nas quais o sujeito é referido como “*One remembers(...), one knows*”. Em português pode optar-se várias formas do sujeito indefinido, embora o efeito produzido possa não ser idêntico ao do texto original. Isto levou a que várias medidas fossem adotadas, de forma a respeitar a mesma forma do TP. Não produzindo, contudo, o mesmo efeito em português:

- “**One remembers** by contrast, the well-known instance in World War I when it took our cryptographic forces weeks or months (...)”. // “Por outro lado, **todos nos lembramos** do famoso exemplo durante a Primeira Guerra Mundial, quando as nossas forças criptográficas demoraram semanas ou meses a determinar que uma mensagem que tinha sido interceptada estava codificada em japonês; e demorou relativamente pouco tempo a decifrá-la, assim que souberam qual era a língua.
- “(...) would of course **lead one to expect** common features in all languages, due to their essentially similar mechanism of development.”  
// “(...) entre outras, **levariam qualquer pessoa** a esperar que todas as línguas tivessem características em comum”

Esta dificuldade foi particularmente complicada de resolver e levou novamente a um repensar das várias opções tradutórias possíveis. Como foi mencionado acima, utilizaram-se possíveis formas do sujeito indefinido, de acordo com a composição do texto de partida. Contudo, a diferença que a utilização deste tipo de sujeito tem,

no português e no inglês respetivamente, produziu efeitos ligeiramente diferentes. Embora as propostas de tradução para o tipo de sujeito utilizado não sejam a melhor equivalência para o texto de partida, respeitam a sua composição e penso que não alteram o carácter argumentativo e expositivo do texto.

### c) Termos Técnicos

Dado que o texto de Weaver é o grande impulsionador e preconizador do conceito de tradução automática, é inevitável esta publicação conter aquilo que são os primeiros termos técnicos sobre o tema. Tanto as especificidades da tradução técnica, como as áreas de estudo que os sistemas incorporam (informática, ciência da computação, língua natural, entre outros) estão na origem de determinados termos específicos, que exigem uma especial atenção. Embora não tenham suscitado grandes dificuldades técnicas, achei pertinente fazer o levantamento de alguns dos termos que obrigaram a uma reflexão mais ponderada. Entre os quais:

1. “Having had considerable exposure to **computer design problems** during the war, and being aware of the speed, capacity, and logical flexibility possible in modern electronic computers, it was very natural for W. W. to think, several years ago, of the possibility that such computers be used for translation.” (Weaver 1949: 14) // “Tendo tido uma exposição considerável aos **problemas de conceção de computadores** durante a guerra, e tendo consciência da velocidade, da capacidade, e da flexibilidade lógica possível dos computadores eletrónicos modernos, é bastante natural que W.W. tenha pensado, há muitos anos atrás, na possibilidade de tais computadores serem utilizados para tradução.”
- Este é um dos muitos exemplos da linguagem dos sistemas de TA. Frequentemente, em muitos textos sobre o tema, encontramos a expressão “computer design”, que na maior parte dos casos refere-se ao *design* (a forma como algo é planeado, desenvolvido e produzido em massa. É uma atividade bastante estratégica, comercial e criativa e tem

objetivos marcadamente de *marketing*; refere-se geralmente ao aspeto exterior de determinado produto) de computadores. No entanto, e lendo cuidadosamente a frase em que “computer design” se insere, podemos perceber que não se refere à configuração física do computador. Esta expressão diz respeito, como foi mencionado na caracterização do texto de partida, ao conhecimento que Weaver adquiriu durante a Segunda Grande Guerra sobre os computadores e de como estes eram concebidos. Conclui-se, por isso, que esta é uma exceção à regra onde o termo “design” se refere à criação ou conceção de computadores, e não à ideia criativa de conceber de forma original o desenho exterior de um computador. É um tipo de expressão que requer rigor, uma vez que tem vários significados – por este motivo, e aqui a leitura cuidada é importante, depois de analisar cuidadosamente o contexto percebe-se que “design” se refere à “conceção” ou “criação” de computadores.

2. “Second – as to the problem of mechanical translation, I frankly am afraid the boundaries of words in different languages are too vague and the emotional and international connotations are too extensive to make any **quasimechanical translation** scheme very hopeful.” (Weaver 1949: 14) // “Em segundo lugar – em relação ao problema da tradução automática, receio francamente que as fronteiras das palavras em línguas diferentes sejam demasiado vagas, e que as conotações emocionais e internacionais sejam demasiado amplas para que um **esquema de tradução quase automático** seja promissor.”

- Este provou ser um dos adjetivos mais complexos e difíceis de resolver em termos de tradução. “Quasimechanical” é um termo que não tem entrada em qualquer dicionário consultado. Existem entradas para “quasi”, que significa “quase: por pouca diferença” (Porto Editora 2009: 854). E existem entradas para “mechanical”, como: “mecânico; técnico” (Porto Editora 2009: 600). Como poderemos observar em vários textos sobre os sistemas de TA, “mechanization” é um termo bastante frequente e que está sistematicamente associado ao processo de tradução, que em português apelidamos de processo de tradução “automática”. Ao contrário do exemplo abaixo, algo “quase mecânico” não faria sentido, uma vez que neste contexto, “mechanical” está a qualificar

a tradução. Por essa razão, optou-se por “uma tradução quase automática”, dado que para um leitor português (tendo ou não conhecimentos na área) o significado seria perceptível (o que não aconteceria com “quase mecânica”).

3. “I will admit that basic English seems to indicate that we can go further than we have generally done in **the mechanization of speech**, but you must remember that in certain respects basic English is the reverse of mechanical (...)” (Weaver 1949: 14) // “Admitirei que o inglês básico parece indicar que podemos ir mais além do que foi feito geralmente **na mecanização do discurso**, mas deve lembrar-se que em certos aspetos, o inglês básico é o oposto do mecânico (...)”
- Imediatamente depois do exemplo acima referido, temos um dos termos mais conhecidos e de frequente utilização no campo da TA. “Mechanization”, deriva do conhecido termo “machine”. Em português, este termo tem muitas vezes como opção tradutória “automático/a”, especialmente quando aliado ao processo de tradução em si. No entanto, se na tradução figurasse algo como “automatização do discurso”, o leitor poderia não saber ao certo o que isso quereria dizer. “Automatização” é de facto um termo informático, contudo, este possui vários significados, entre eles: “técnica de que resulta a execução automática de tarefas administrativas, científicas ou industriais; uso de aparelhos eletrónicos para controlo de funcionamento de um sistema ou processo (Infopedia 2013). Por esta razão, decidi-me por “mecanização”. Não só por ser a correspondência imediata do termo original para “mechanization”, mas também porque estudando atentamente o significado de “mecanização”, este significa algo que é “efetuado por meio de uma máquina; tornar maquinal” – que exprime exatamente o que tem sido feito ao discurso.
4. “This may be, for all I know, a famous theorem of philology. Indeed the well-known **bow-wow, woofwoof, etc. theories** of Müller and others, for the origin of languages (...)” (Weaver 1949: 13) // “Este pode ser, tanto quanto sei, um famoso teorema de filologia. Na verdade, as conhecidas teorias de Müller e

outros para a origem da linguagem, como a **teoria *bow-wow*, *woof-woof***, entre outras, levariam qualquer pessoa a esperar que todas as línguas tivessem características em comum (...)

▪ Embora não seja relativo a uma terminologia do foro informático e/ou dos sistemas de TA, é um exemplo do quão rigoroso um texto técnico é, e de quão necessária a extração da informação principal do texto é para que o leitor não se perca aquando da leitura. A “teoria bow-wow” é uma das muitas teorias existentes relativas à origem da linguagem, que tem como autor Max Müller. Müller<sup>10</sup> fora um filólogo conhecido, que sempre demonstrara interesse em estudos comparatistas, e que acreditava que a origem da linguagem estava na imitação de sons naturais – aquilo a que chamamos onomatopeias. Müller acreditava que os humanos tinham desenvolvido a linguagem através da reprodução de sons naturais: da água, do vento, dos animais, entre outros, associando o som a coisas e, assim, expressando o que queriam dizer. “Bow-wow” é, como exemplo disso, uma onomatopeia inglesa que expressa o latido de um cão. Inicialmente, figurava na tradução algo como “as onomatopeias “*bow-wow* [ão-ão] e *woof-woof* [ão-ão]”, uma vez que ambas as teorias têm o nome de onomatopeias, exatamente porque se referem às mesmas para explicar a forma como a linguagem surgiu, para nós, humanos. No entanto, embora sejam onomatopeias, estas são acima de tudo teorias que apenas assumiram como nome a ideia que as preconizou. Por esta razão, figura na tradução “as conhecidas teorias de Müller e outros para a origem da linguagem, como a teoria do *bow-wow*, *woof-woof*<sup>11</sup>, entre outras (...)”.

5. “This theorem states that a robot (or a computer) constructed with **regenerative loops** of a certain formal character is capable of deducing any legitimate conclusion from a finite set of premises.” (Weaver 1949: 16) // “Este teorema

---

<sup>10</sup> [http://www.psych.yorku.ca/gigi/documents/Thorndike\\_1943.pdf](http://www.psych.yorku.ca/gigi/documents/Thorndike_1943.pdf)

<sup>11</sup> Esta opção tradutória pode parecer estranha para o leitor português, uma vez que o nome desta teoria não é traduzido. Isto porque a denominação da teoria de Müller tem por base duas onomatopeias (*bow-wow*, *woof-woof*) que designam o mesmo som: o ladrar de um cão, em português “ão-ão”. Depreende-se, depois da investigação feita e argumentada acima, que a utilização destas onomatopeias para o mesmo som está, também, na origem das línguas. Duas maneiras diferentes de reproduzir o mesmo som pode explicar como é que surgiram duas línguas diferentes, por exemplo. Dado que em português existe apenas uma onomatopeia para o ladrar de um cão: “ão-ão”, achei que os princípios por detrás da escolha das onomatopeias de Müller se perderiam e, por conseguinte, optei por manter o nome original.

declara que um robô (ou um computador) construído com **loops regenerativos** de um certo carácter formal é capaz de deduzir qualquer conclusão legítima de um número finito de premissas.”

- Inicialmente, este foi um dos termos que não reconheci imediatamente como “técnico”, uma vez que o substantivo “loop” tem vários equivalentes em português. Dado o contexto, percebe-se que um computador, se construído com “loops” – espirais; ciclo (vicioso) de um programa e/ou comportamento do mesmo que se repete vezes sem conta, de forma interminável – após um número de tentativas, consegue concluir qual a melhor escolha. No entanto, na área da informática este é dos muitos termos que, dada a falta de equivalência de uma palavra que exprimisse exatamente o seu significado em português, é intraduzível. Segundo especialistas que trabalham no terreno, quando algum programa ou até mesmo o próprio computador possui alguma inconsistência que é sequencial, diz-se que o mesmo está em “loop”. Na tradução figurou, por conseguinte, o estrangeirismo e não a opção “espirais” que originalmente tinha sido considerada.

6. “Now there are surely **alogical elements** in language (intuitive sense of style, emotional content, etc.) so that again one must be pessimistic about the problem of *literary* translation.” (Weaver 1949: 16) // “Hoje em dia existem certamente **elementos ilógicos** na linguagem (sentido de estilo intuitivo, conteúdo emocional, entre outros) para que tenhamos que ser pessimistas acerca do problema de tradução *literária*.”

- Embora pareça um exemplo de tradução relativamente simples, após alguma pesquisa demonstrou que implicava mais cuidado e atenção do que se supunha. Pesquisando no Dicionário da Porto Editora, edição de 2009, não existe qualquer entrada para “alogical”. É apenas *online* que encontro definições para este termo. “Alogical” significa portanto, algo “outside the bounds of that to which logic can apply.” (Dictionary 2012). O oposto de “lógico” ou de algo “razoável”. A mesma pesquisa denota que na língua inglesa “alogical” é utilizado para fazer referência a objetos concretos, enquanto “illogical” se refere a algo mais abstrato. Inicialmente, a tradução

que figura era exatamente “alógicos”, uma vez que em português na área da filosofia, “alógico” significa exatamente algo “que não é regido pelos princípios nem pelas regras da lógica” (Infopedia 2012). No entanto, as primeiras entradas dão conta de que “alógico” é algo que é evidente; algo que não precisa de provas para se demonstrar correto. Por este motivo, o adjetivo escolhido foi ‘ilógico’, por ser o antónimo de “lógico”, ou seja, algo contrário à lógica.

7. “Claude Shannon, of the Bell Telephone Laboratories, has recently published some remarkable work in the **mathematical theory of communication**.” (Weaver 1949: 16) // “Claude Shannon, dos laboratórios *Bell Telephone*, publicou recentemente alguns trabalhos notáveis sobre a **teoria matemática da comunicação**.”

- Não é uma opção tradutória que levante problemas. No entanto, e de forma a elucidar o leitor, é importante sublinhar que esta “teoria matemática da comunicação” é uma das publicações mais importantes do trabalho de Claude Shannon. Sem esta teoria, Weaver não tinha um suporte sólido para o princípio de descodificação que mais tarde daria lugar aos ideais de TA. A Teoria da Informação consistia num “sistema geral de comunicação” (Infoescola 2013) composto por elementos específicos que iam desde a origem da informação, passando pelo conteúdo e codificação, até à posterior descodificação para chegar ao destinatário. Foram estas características que suscitaram o interesse de Weaver, e que este ajudou a desenvolver. Foi este o suporte matemático que deu vida aos ideais de Warren Weaver.

8. “During the war Shannon wrote a most important analysis of the whole cryptographic problem, and this work is, W. W. believes, also to appear soon, it having been **declassified**.” (Weaver 1949: 16). // “Durante a guerra, Shannon fez uma análise muito importante de todo o problema criptográfico, e este trabalho como acredita W. W., está para ser publicado brevemente, uma vez **desclassificado**.”

- Este é um exemplo de uma expressão aparentemente simples e que, no entanto, levanta mais questões do que se suporia. Numa primeira leitura, “declassified” pareceu-me ser um “false friend”, uma vez que é bastante semelhante com o adjetivo “desclassificado” em português. Por esta razão, na tradução passou a figurar algo como “assim que a sua confidencialidade fosse levantada”. Tendo em conta o contexto da frase e sabendo que “desclassificado” se refere a “algo desacreditado; que não obteve consideração” (Infopedia 2013) presumiu-se que não seria uma boa escolha, e que a alternativa encontrada era a melhor opção. Observando cuidadosamente o contexto, pode-se concluir que o que Warren Weaver pretendia dizer é que havia um determinado documento que seria publicado, assim que deixasse de ser confidencial. “Declassified” tem, portanto, um forte cariz político e prende-se com todo o tipo de documentos e/ou informações sigilosas, na posse do Estado e seus organismos. Exemplo disso é o seu conhecido antónimo “classified”, normalmente acompanhado do nome “information” – designando exatamente um tipo de informações secretas que não podem ser reveladas (também cunhado com a designação “top secret”). Em Portugal, o termo “desclassificado” tem uma conotação jurídica, referente à decisão do juiz em mudar a classificação do crime cometido pelo réu e pela qual este é acusado (pt.scribd 2013). No entanto, segundo uma fonte fidedigna, “desclassificado” começa a ser também um termo que, tal como acontece no inglês, começa a ser utilizado para designar documentos cuja confidencialidade cessou. Como pequeno exemplo, a ANA aeroportos© tem documentos *online* que falam precisamente sobre a gestão e controlo de documentos oficiais, nos quais o termo é utilizado com o mesmo significado. Segundo o documento e passo a citar: “A desclassificação tem lugar logo que cesse a causa que determinou a classificação do documento ou quando a classificação tiver sido incorretamente atribuída” (Dgtf 2013) – demonstrando que o termo, aparentemente enganador, compreende e classifica o mesmo tipo de informação, que uma vez levantada a confidencialidade, é “desclassificada”.

9. “And it is one of the chief purposes of this memorandum to emphasize that **statistical semantic** studies should be undertaken, as a necessary preliminary step.” (Weaver 1949: 16) // “É um dos objetivos principais deste memorando



ênfatisar que se devem fazer os estudos da *semântica estatística*, como uma etapa preliminar necessária.”

▪ Este é o último exemplo dos termos técnicos e provavelmente o mais complexo. A expressão “statistical semantic” é um termo que o próprio Warren Weaver utilizou pela primeira vez nesta publicação. Está intimamente ligada ao campo da língua natural, e diz respeito à desambiguação efetuada pelos sistemas através da quantidade de vezes que determinadas palavras ocorrem. Em português, dentro dos tipos de semântica que conhecemos, é um termo que não existe. No entanto, alguns estudos sobre língua natural incluem o termo traduzido que figura na tradução<sup>12</sup>. Por ser uma área de estudo intimamente ligada com a linguística e a ciência computacional, é um tipo de semântica que é composto por uma componente estatística. Por esta razão, e porque não é algo completamente estranho para o leitor, optou-se por uma tradução literal do termo.

---

<sup>12</sup> Como é exemplo, o planeamento do seminário de “Língua Natural”, no instituto Superior Técnico. Ver em: <https://fenix.ist.utl.pt/disciplinas/ln-2/2012-2013/1-semester/planeamento>, presente nas referências bibliográficas.

### **3. Aplicação ao modelo de análise textual ao texto *The Development and use of machine translation systems and computer-based translation tools*, de John Hutchins (1999)**

#### **3. 1. Análise dos fatores extratextuais**

William John Hutchins (AUTOR) nasce a 27 de Janeiro de 1939, na cidade de Londres. Filho de William Arthur Hutchins (1907-1972) e Frances Helen Read (1913-2001), John Hutchins começa o seu notável percurso em 1960 quando adquire o bacharelato em línguas (francês e alemão) na Universidade de Nottingham. Dois anos mais tarde obtém um diploma de bibliotecário na Universidade de Londres, tendo desempenhado essa função até 1965 na Universidade de Durham. Muda-se, depois, para Sheffield (onde, de resto, se casa e constitui família) nesse mesmo ano, para passar a ser auxiliar de bibliotecário na Universidade de Sheffield até 1971.

Os anos mais importantes da vida profissional de Hutchins começam quando, neste ano, se muda para a Universidade de East Anglia e se torna auxiliar sénior de bibliotecário, função desempenhada até 1980. É neste intervalo de tempo que o autor escreve vários artigos e livros sobre linguística, como *Languages of indexing and classification. A linguistic study of structures and functions* (1975). Porém, é só a partir de 1978<sup>13</sup> que John Hutchins se debruça sobre a área de TA e começa a dedicar artigos e obras a esta temática. Por este motivo, o intervalo de tempo entre 1980 e 1998, em que Hutchins se torna sub-bibliotecário na mesma Universidade, é particularmente relevante. Em 1986, Hutchins compõe aquela que é uma das obras mais importantes da TA: *Machine translation: past, present, future*, que veio esclarecer e elucidar muitos alunos, linguistas, tradutores e até cientistas sobre o conceito, a história e as lacunas da TA (INTENÇÃO, FUNÇÃO DO TEXTO e PÚBLICO). *Introduction to machine translation* (1992), que escreve juntamente com Harold L. Somers, é outra importante obra da sua bibliografia que fala sobre as características fundamentais dos sistemas de

---

<sup>13</sup> In <http://www.hutchinsweb.me.uk/biography.pdf> (2013)

TA e quais as bases linguísticas em que assentavam. A contribuição extraordinária que fez até então para a área da TA, possibilitou que em 1992 se tornasse editor do *MT News International*, um boletim informativo da Associação Internacional de TA, função que desempenhou até 1997. Para além deste importante marco profissional, Hutchins tornou-se também presidente da Associação Europeia de TA (1995-2004) e da Associação Internacional de TA (1999-2001).

Distinguido em 2001 com o prémio IAMT de honra, Hutchins organizou ainda o compêndio de *software* de tradução desde 2000 e o Arquivo de TA, desde 2004. A última contribuição de Hutchins conhecida até à data é o artigo “Machine translation: history of research and use” in *Encyclopedia of Languages and Linguistics*, de 2006<sup>14</sup>.

### ***3.1.1. The Development and use of machine translation systems and computer-based translation tools***

O artigo *The Development and use of machine translation systems and computer-based translation tools* foi publicado por John Hutchins na Conferência Internacional de TA e Processamento de Linguagem de Informação Computacional (MEIO), que teve lugar na cidade de Pequim, China (ESPAÇO) entre 26 e 28 de Junho, de 1999 (TEMPO).

Através desta publicação Hutchins não só faz uma retrospectiva geral sobre a história dos sistemas, como fala sobre a produção de documentação técnica e a importância desta para os sistemas. O autor pretende, através desta breve análise, elaborar sobre quais as ferramentas de *software* mais importantes e qual a dimensão que a TA tem adquirido com a chegada da *internet*, em comparação com a tradução humana (MOTIVO, FUNÇÃO DO TEXTO e INTENÇÃO)

---

<sup>14</sup> In <http://www.hutchinsweb.me.uk/history.htm>

### **3.1.2. Reflexo dos fatores extratextuais no texto**

O autor pretende, através da formulação deste texto, informar o leitor (INTENÇÃO) sobre a área da TA: em que consiste, como surgiu, que programas foram desenvolvidos e em que países, quais os mais e menos relevantes, qual a importância das ferramentas de *software*, quais os desenvolvimentos que advieram com a chegada da *internet* e qual o lugar da TA em comparação com a tradução humana. Avaliando a quantidade de temas que aborda e desenvolve, percebe-se que Hutchins pretende tornar disponível esta informação a um público que se pressupõe ter conhecimentos elementares de informática, e poucos sobre TA em geral, colmatando a lacuna de informação que existe sobre a área (MOTIVO). O texto tem como objetivo principal fornecer informações (FUNÇÃO DO TEXTO).

Dado que estamos perante um texto onde o conhecimento é transmitido de forma objetiva e precisa, e dado que o texto é composto por muita terminologia técnica, podemos dizer este texto tem um cariz menos científico que o anteriormente analisado. Como sublinha Fontanet, “científico e técnico” são termos que são frequentemente agrupados como se designassem apenas um conceito. Contudo, a autora faz uma distinção sucinta, na qual afirma que o conteúdo de um texto técnico não envolve emoções e não recorre à argumentação para persuadir o seu leitor, como o texto científico faz (Fontanet 2011: 3). O objetivo do texto técnico é ser funcional, transmitir diretamente a realidade técnica que o compõe, sem induzir o leitor em erro ou confusão. É por este motivo que se considera que o segundo texto, agora em análise, embora tenha características em comum com o texto de Weaver, possui um cariz mais técnico que o seu antecedente. As referências que o autor faz a datas, países e nomes de empresas sustentam a validade da informação apresentada são características que ambos os textos apresentam (COMPOSIÇÃO E CONTEÚDO).

No entanto, são características de cariz mais texto técnico:

- a) Abundância de termos técnicos como: *cross-language information retrieval*, *custom-built MT*, *components of information access systems*,

*localisation software, MT for interchange, Multilingual Word Processing, Spoken Language MT*, entre outros.

- b) Referências a países, empresas, datas e autores:
- i. “Until the late 1980s, systems of all these kinds were developed, and it is true to say that all current commercially available systems are also classifiable into these three basic system types: direct, interlingual and ‘transfer’.” (Hutchins 1999: 3).
  - ii. “But the best known systems adopting the ‘transfer’ approach were research projects: Ariane at GETA in Grenoble, an MT project going back to the 1960s, and Eurotra funded by the Commission of the European Communities.” (Hutchins 1999: 3).
  - iii. “While the translation of many legal texts continues to be done by human translators, the Systran systems are used increasingly not only for the translation of internal documents (with or without post-editing) but also as rough versions for the assistance of administrators when composing texts in non-native languages (Senez 1996).” (Hutchins 1999: 4).
  - iv. “Also using Logos are Ericsson, Osram, Océ Technologies, SAP and Corel. Systran has many large clients: Ford, General Motors, Aérospatiale, Berlitz, Xerox, etc. The METAL German-English system has been successfully used at a number of European companies: Boehringer Ingelheim, SAP, Philips, and the Union Bank of Switzerland.” (Hutchins 1999: 5).

### **3.2. Análise dos fatores intratextuais**

#### **a) Assunto**

De acordo com a convenção dos textos científicos e técnicos, o assunto principal é mencionado no título da publicação. O assunto principal é mencionado no título da publicação, bastante claro e objetivo. Todos os elementos isotópicos do texto confirmam isso mesmo. No entanto, é importante referir que antes de aprofundar a temática que pretende, Hutchins debruça-se sobre os tipos de procura de tradução, fazendo um breve percurso histórico. O autor finaliza com breves notas sobre a

importância da *internet* na área da TA, de forma a demonstrar o quão importante foi para a expansão dos sistemas que foram aparecendo, disponibilizando serviços de tradução em rede e permitindo uma maior utilização de ferramentas de *software*. O último capítulo do artigo de Hutchins é uma breve reflexão, onde o autor compara a tradução humana à automática, e afirma que embora a TA tenha vindo a ganhar destaque, existem tarefas que só um tradutor humano pode desempenhar.

## **b) Conteúdo**

A análise do conteúdo demonstra, ao mesmo tempo, a macroestrutura do texto:

1. **Abstract (e respetivas *keywords*):** O autor anuncia que assuntos vai tratar, nomeadamente, o *software* de tradução que é determinante para traduções assistidas por computador que possam ser publicadas, os desenvolvimentos para sistemas de linguagem controlada, o *software* de localização e de trabalho do tradutor. O autor fala ainda sobre os desenvolvimentos de *software* de tradução para utilizadores não-profissionais, especialmente através de páginas *Web* e outras aplicações *online*.
2. **Types of Translation demand:** Hutchins começa por falar nos tipos de procura de tradução que deram origem aos sistemas e ferramentas historicamente relevantes: a disseminação, a assimilação, TA para intercâmbio e componentes de sistemas de acesso a informação. São estes os conceitos que o autor diz serem fulcrais para o desenvolvimento e uso dos sistemas de TA que conhecemos hoje. A disseminação e a assimilação dizem respeito ao nível de qualidade de uma tradução efetuada por um sistema e a interesses e finalidades completamente distintas entre si aquando da opção por uma tradução automática. Enquanto a disseminação se refere à procura de um tipo de tradução com o mesmo nível de qualidade de um tradutor humano (embora a tradução tenha sempre que ser corrigida, posteriormente, por um tradutor humano), a assimilação tem a ver com um tipo de tradução de qualidade inferior, sendo esta preterida desde que o utilizador compreenda o conteúdo principal. Por outro lado, os restantes tipos de tradução têm propósitos ligeiramente diferentes. O terceiro tipo de procura de tradução, a TA para intercâmbio, é relativa à tradução imediata de um tipo de comunicação feito exclusivamente por meios tecnológicos: textos de páginas da

*Web, emails*, entre outros. Por fim, o último tipo trata a tradução dentro de sistemas multilingues para extração (ou recuperação) de informação. Diz Hutchins que são estes quatro tipos de procura que servem as finalidades de programação de sistemas vindouros.

3. **Historical background:** De modo a que o leitor entenda um pouco mais como surgiram os primeiros sistemas, o autor faz um breve percurso histórico. Neste subcapítulo Hutchins fala da importância do computador eletrónico, e de como foi a partir da invenção deste nos anos 40 do século XX que se pensou em traduzir textos em algumas línguas. Numa primeira fase, os sistemas faziam traduções diretas a partir de dicionários bilingues. Contudo, a partir dos anos 80 do mesmo século, avanços na linguística computacional permitiram que se tentasse uma abordagem indireta. Esta abordagem consistia na análise da língua-fonte através de representações abstratas de “significação” (i.e. diretrizes que identificavam a morfologia, a sintaxe e semântica para resolver o problema recorrente da ambiguidade). Estas representações abstratas tinham como objetivo não só resolver os problemas de ambiguidade que surgissem, mas também possibilitar que vários textos pudessem ser traduzidos para uma ou mais línguas-alvo, razão pela qual a abordagem por interlíngua adquiriu uma dimensão importante. É por este motivo que nos dias de hoje todos os sistemas comerciais que se encontram disponíveis podem ser classificados de acordo com um dos seguintes tipos de sistema: direto, indireto e por transferência. O autor faculta exemplos de sistemas atuais para cada um destes tipos.
4. **Governmental and non-commercial use:** O autor faz uma breve menção ao uso governamental e não comercial dos sistemas, lembrando que as primeiras instalações e serviços advieram de interesses militares e políticos – áreas que podiam suportar os custos elevados do *hardware* necessário. Hutchins destaca a Força Aérea Americana, que introduziu o sistema Systran em 1970 para conseguir traduzir documentação militar e científica russa. O Centro Nacional de Inteligência Espacial assumiu o controlo dessas operações, e produz agora traduções para várias organizações governamentais americanas. Na Europa, o maior serviço de tradução é o da Comissão Europeia, que foi uma das primeiras organizações a instalar a TA, em 1976.

5. **Production of technical documentation:** Este ponto é importante para falar sobre a importância da documentação técnica, e da assunção que se tinha nos anos 90 (do século XX) de que o intuito dos sistemas de tradução automática era o de traduzir documentação técnica com uma qualidade que pudesse ser publicada. No entanto, cedo se percebeu que os sistemas não conseguiam substituir o tradutor humano e as aptidões que possui. As traduções automáticas continuariam a ter procura por terem um custo mais baixo que o emprego de um ou mais tradutores humanos, contudo, a tradução humana continuaria a ser necessária para a revisão. Este é um dos pontos principais dos sistemas, sublinha Hutchins. Esta funcionalidade semi-eficiente dos sistemas continua a ser a ideal para as grandes empresas, e para a documentação técnica que é pedida em grande escala. O desempenho de determinado sistema dentro de uma grande empresa é um dos pontos fulcrais de análise para o autor, e é aqui que entram muitas das ferramentas de *software* de TA, fundamentais para auxiliar tanto o sistema, como a vindoura tradução.
6. **Controlled language and domain-specific systems:** Após analisar a importância da documentação técnica, observou-se que havia a necessidade de existir um certo nível de controlo, para que os problemas de ambiguidade fossem menores e a produção de documentação fosse maior e mais rápida. Desta necessidade crescente em diminuir as incongruências dos sistemas e, consequentemente, aumentar as traduções de documentação técnica que contêm uma terminologia, por vezes própria de uma determinada empresa para exprimir conteúdo e marca, nasceram os sistemas de linguagem controlada. Estes sistemas foram concebidos pelas próprias empresas, servindo os seus interesses e dos seus clientes. Cedo muitas mais empresas desenvolveram o seu próprio *software*, e com mais línguas disponíveis. É aqui que as ferramentas de *software* são cruciais para o funcionamento dos sistemas e para a produção em larga-escala da TA, sublinha o autor. Hutchins lembra três ferramentas de extrema importância, que não só ainda existem no mercado, como são bastante importantes na área da TA: a memória de tradução, o *software* de trabalho do tradutor e o *software* de localização.
7. **Translation workstation:** O autor analisa a importância desta ferramenta de tradução automática, descrevendo em que consiste e como pode ser utilizada. O



*software* de trabalho do tradutor nasceu da necessidade de auxiliar o tradutor na revisão e processamento da documentação técnica, especialmente aquela que era específica de uma determinada empresa. Este *software* compila várias ações e serviços, nomeadamente o processamento de palavras, dicionário de sinónimos, contagem de palavras e até memória de tradução integrada. Contudo, notou-se que a última ferramenta, por si só, já desempenhava operações de grande utilidade e importância, entre as quais o armazenamento dos textos originais e as suas versões traduzidas, alinhando-as lado-a-lado. A relevância deste *software* recai exatamente no armazenamento e memória de expressões, ou frases inteiras, de ambos os textos de trabalho. Isto permite que muitas empresas se submetam a revisões de muita da sua documentação técnica para atualizar o seu arquivo. O tradutor pode, com esta ferramenta, localizar especificamente o excerto do texto que pretende tendo, simultaneamente, acesso à terminologia. Tanto este *software*, como o de trabalho do tradutor vieram revolucionar o mundo da TA, não só pela inovação e funcionalidade que vieram acrescentar aos sistemas, como também pelo poder que deram ao tradutor humano de ter o controlo de todo o processo tradutório.

8. **Localisation of software:** Com abordagens semelhantes, existe aquela que é uma das ferramentas mais procuradas na área da TA: o *software* de localização. Embora não consista nas mesmas operações que as ferramentas descritas anteriormente, a sua inclusão não é feita ao acaso e, como relembra Hutchins, sem este *software* muitos dos textos técnicos não estariam disponíveis na quantidade de línguas que hoje é possível encontrar. Esta é a ferramenta que protege o código do programa ao permitir a extracção do texto a ser traduzido num ambiente seguro, no qual a versão traduzida pode ser inserida no local certo e no ficheiro adequado (Hartley *apud* Baker 2006: 120).
9. **Systems for personal computers:** Neste subcapítulo o autor sublinha a importância comercial que o *software* de tradução para computadores pessoais assumiu, especialmente nas empresas japonesas nos anos 80 do século XX. Fora do Japão, os sistemas para computadores pessoais apareceram relativamente mais cedo, embora em poucas empresas. Hutchins sublinha os sistemas ALPS, concebidos para aceder e criar recursos terminológicos. A importância deste tipo de sistemas recai no facto de permitir que qualquer utilizador, seja qual for o seu

*background*, execute traduções (ora *freelancer*, ora apenas com o intuito de querer comunicar noutras línguas) sem se incomodar com a qualidade da mesma. No entanto, foi só no final dos anos 80 do século XX que surgiram os sistemas mais comerciais e que se encontram ainda no mercado, onde se destaca a Globalink. A inclusão destes sistemas na publicação de Hutchins é essencial para demonstrar o quão têm contribuído para a venda de *software* de tradução, e para fazer a ponte entre a expansão da utilização destes serviços e a chegada da *internet*.

10. **MT on the internet:** A invenção desta rede global vem influenciar ainda mais os já existentes sistemas de TA para computadores de uso pessoal, e revolucionar a área da tradução ao disponibilizar serviços em rede para qualquer utilizador comum. O domínio global da *Web* veio fornecer serviços de tradução em rede, expandindo a tradução em grande escala para clientes de grandes empresas nesta plataforma. A influência direta e a importância da *internet* na área da TA faz-se notar não só no crescente número de *software* de tradução para páginas *Web*, como na tradução *online* de inúmeras línguas. Embora ainda apresente algumas lacunas é uma área que merece bastante destaque, não só porque permite que outras opções para os sistemas sejam pensadas e enquadradas, como o cariz universal de uma plataforma como a *internet* vem interessar cada vez mais qualquer utilizador, profissional ou amador, bem como as grandes empresas.
11. **Future needs and developments:** Apesar de todas as atualizações e crescimento recente nas áreas que o autor descreve, existe ainda a necessidade de haver uma tradução de documentos para línguas estrangeiras menos utilizadas, que sejam de baixo custo. Sendo que o custo, sublinha Hutchins, é a palavra-chave para os objetivos futuros, especialmente a nível de *software*. O autor lembra ainda neste subcapítulo que a maior ambição da TA é conseguir a tradução de língua falada. Contudo, trata-se de um objetivo muito dispendioso e que ainda requer investigação.
12. **Comparison of human and machine translation:** Ao fazer o sumário de tudo o que compilou e descreveu, Hutchins faz questão de deixar bem claro que existem grandes diferenças entre a tradução humana e a automática. Porém, o autor lembra que a comparação entre as duas tem que ser feita a nível de funcionalidade e de procura. Se o objetivo é um tipo de procura que diz respeito

à disseminação, ou seja, a tradução de textos com uma qualidade que possa ser publicável, então a tradução humana é a mais satisfatória. A TA perde por ser mais dispendiosa que um tradutor humano, uma vez que exige a manutenção de *software* e ainda o custo da revisão. Contudo, se o tipo de procura for o de assimilação então a TA fica a ganhar, uma vez que os tradutores humanos não estão preparados, e até equipados, para fazer uma tradução “rascunho” de um texto técnico, que é executada apenas para ser vista por alguém que quer ter uma noção do conteúdo principal. Estas são as conclusões principais com as quais Hutchins encerra a sua publicação, após ter elaborado sobre os sistemas e ferramentas mais relevantes e qual a sua difusão.

13. **References:** Referências bibliográficas, organizadas alfabeticamente.

**c) Pressuposição**

Ao contrário do texto de Warren Weaver anteriormente analisado, na publicação de Hutchins não existem elementos textuais pelos quais se infira que o leitor já sabe do assunto, ou que o autor se dirija a ele como se ele tivesse conhecimentos proficientes dos temas explorados. É por este motivo que este artigo contém subcapítulos (2, 3, 6, 7, 8 e 10 dos descritos na secção acima) onde o autor descreve a informação apresentada, dizendo em que consiste, qual a sua relevância e repercussões no mercado ou na área da TA. Isto pressupõe que o autor do texto considerou a hipótese do seu texto ser lido por pessoas menos instruídas ou conhecedoras desta área.

No entanto, é importante sublinhar que embora Hutchins simplifique muita da informação, ele pressupõe ao mesmo tempo que o leitor tem alguns conhecimentos de informática e/ou tecnologias aplicadas à tradução. Isto porque embora ele faça um breve percurso histórico e explique o que despoletou a necessidade e o desejo de tornar a TA possível, o autor nunca explícita em que consiste um *software*, qual a diferença entre *software* e *hardware* ou ainda como criar um sistema de TA operacional, por exemplo. Isto leva a crer que o autor parte do princípio que mesmo que o seu leitor não saiba o suficiente sobre TA, saberá com certeza o suficiente sobre informática ou tecnologia aplicada à tradução para perceber a informação que vai transmitindo. São exemplos:

- i. “With the third type – MT for **interchange** – the situation is changing quickly. The demand for translations of electronic texts on the Internet, such as Web pages, electronic mail and even electronic ‘chat’ lists, is developing rapidly. In this context, the possibility of human translation is out of the question. The need is for immediate translation in order to convey the basic content of messages, however poor the input. MT systems are finding a ‘natural’ role, since they can operate virtually or in fact in real-time and on-line and there has been little objection to the inevitable poor quality. Another context for MT in personal interchange is the focus of much research. This is the development of systems for spoken language translation, e.g. in telephone conversations and in business negotiations.” (Hutchins 1999: 2)
- ii. “The fourth type of MT application – as components of information access systems – is the integration of translation software into: (i) systems for the search and retrieval of full texts of documents from databases (generally electronic versions of journal articles in science, medicine and technology), or for the retrieval of bibliographic information; (ii) systems for extracting information (e.g. product details) from texts, in particular from newspaper reports; (iii) systems for summarising texts; and (iv) systems for interrogating non-textual databases. This field is the focus of a number of projects in Europe at the present time, which have the aim of widening access for all members of the European Union to sources of data and information whatever the source language.” (Hutchins 1999: 2)
- iii. “Until the 1990s the normal assumption was that MT systems were intended to be used for the production of documentation of publishable quality, primarily but not exclusively of a scientific and technical nature. The assumption was, in other words, that MT systems were to be used in conditions where otherwise human translators would be employed with expertise in the subjects concerned. Evidently, the actual quality of MT output was inadequate for direct use. It had to be extensively revised before it could be published, and translators were therefore employed as ‘posteditors’.” (Hutchins 1999: 4)
- iv. “However, few experienced observers expect dramatic developments in this area in the near future – the development of MT for written language has taken

many years to reach the present stage of widespread practical use in multinational companies, a wide range of PC based products of variable quality and application, growing use on networks and for electronic mail.” (Hutchins 1999: 13)

#### **d) Composição**

O texto de Hutchins é composto por onze subdivisões, onde a temática apresenta uma progressão cronológica. De modo a dar ênfase aos sistemas de TA mais importantes comercialmente, o autor remonta à invenção do computador eletrônico em 1940, para explicar como surgiu a TA. A partir daí, Hutchins vai avançando no tempo, relatando as primeiras abordagens e os primeiros sistemas a surgirem, tanto a nível comercial, como para uso pessoal, até chegar ao tema que pretende: os principais sistemas e ferramentas de TA. O autor acaba o seu texto situado no final dos anos 90 do século XX, falando das empresas que se destacaram pelo uso ou venda dos sistemas e ferramentas que mencionou. A data mais recente (esta publicação é de 1999) é o ano de 1998, que aparece justamente no subcapítulo intitulado “Necessidades e desenvolvimentos futuros”.

O texto apresenta esta estrutura para que o leitor consiga acompanhar a progressão lógica dos temas que o autor aborda, e de como se encontram interligados. É um texto que fornece informações, explicações e algumas enumerações de forma prática, onde Hutchins esclarece qual o funcionamento e vantagens de muitos dos sistemas e ferramentas da área da TA, explicando de forma clara porque tiveram sucesso ou não. Não há qualquer desvio dos temas que o autor pretende retratar, e este cariz técnico do texto é reforçado pela abundância e variedade de termos técnicos usados, típicos de uma “língua especializada”, como lembra Mathilde Fontanet (Fontanet 2011: 2). Estamos, porém, não só perante um texto composto por um léxico terminológico rico, mas também por várias referências, datas, nomes e factos.

#### **e) Elementos não-verbais**

O primeiro elemento não-verbal que salta à vista é o espaço que existe entre cada subdivisão do texto. A razão para este *layout* prende-se com a sua composição, e

reside na evidente mudança de tema. É ainda importante salientar que a utilização do espaçamento referido confere uma sensação de ordem, que ajuda à memória do leitor, facilitando a leitura e a compreensão do texto.

#### **f) Léxico**

Uma vez que estamos perante um texto técnico, uma das características mais importantes a destacar é o léxico. Por ser um texto objetivamente sobre e para a área da TA, esta publicação é composta por uma extensa lista terminológica, característica da área dos sistemas e das tecnologias aplicadas à tradução. Ao fazer uma breve retrospectiva sobre os sistemas até chegar à temática que pretende, o autor faz uso de vários termos e/ou expressões que não pertencem ao léxico da língua portuguesa. Falamos, portanto, de empréstimos neológicos de língua inglesa cuja entrada no português nas propostas de tradução apresentadas, “dada a posição ocupada atualmente por Portugal e pelos países lusófonos no contexto internacional (...) é (...) inevitável.” (Correia 2005: 2). Os termos que podemos encontrar no texto são conhecidos e utilizados na língua de chegada como empréstimos ou estrangeirismos, dado que o país (os E.U.A.) e a língua (inglês) são proeminentes no uso.

É importante sublinhar que a riqueza lexical do texto é igualmente uma marca da sua composição. Existem ainda marcas lexicais que permitem perceber que há um sentido histórico desde o início do texto até ao final, e algumas referências que nos permitem pressupor qual o tempo de produção do texto (ignorando o cabeçalho que indica onde a publicação de Hutchins foi apresentada). São exemplos:

- i. “Among the most important of the mainframe ‘transfer’ systems was METAL, supported **for most of the 1980s** by Siemens in Germany. However, it was only **at the end of the decade** that METAL came into the market, and sales were poor. **During the 1990s**, rights to METAL have been transferred to two organisations (GMS and LANT) in a complex arrangement.” (Hutchins 1999: 3)
- ii. “These highly successful systems (**now also available** to users outside PAHO) are general-purpose systems, not constrained in vocabulary or text

type, although obviously the dictionaries are strongest in the health-related social science fields (**Leon and Aymerich 1997**).

- iii. **In the 1990s there have been a number of other examples.** In Finland, the Kielikone system was developed originally as a workstation for Nokia Telecommunications. Subsequently, versions were installed at other Finnish companies and the system **is now being marketed** more widely (**Arnola 1996**).” (Hutchins 1999: 6)
- iv. **“There are now** four main vendors of workstations: Trados (probably the most successful), STAR AG in Germany (Transit), IBM (the TranslationManager), and LANT in Belgium (the Eurolang Optimizer, previously sold by SITE in France).” (Hutchins 1999: 7)

**São termos técnicos predominantes no texto:**

- Controlled Language Systems
- Computer-aided translation
- Computer-based translation
- Computacional linguistics
- Chat rooms
- Corpus-based methods
- Cross-Language Retrieval
- Custom-built MT
- Knowledge-based MT
- Localisation software
- Multilingual word processing
- MT for interchange
- Speech Recognition
- Spoken Language MT
- Sublanguage systems
- Translation/Translator Workstation (or Workbench)
- Translation Memory
- Windows-based versions

**g) Estrutura das frases**

O texto é composto por vários parágrafos, geralmente um pouco longos. Como já foi referido, a existência desta quantidade de parágrafos assegura a atenção do leitor ao indicar que haverá uma provável mudança de tema, ou que se dará destaque a outro tipo de informação secundária. Como é característico dos textos técnicos, as estruturas sintáticas são geralmente um pouco mais longas que o habitual neste tipo de texto. São exemplos disso:

- “This combines access to the Commission’s own very large multilingual database (Eurodicautom), the dictionary resources of Systran, facilities for individual and group terminology database creation and maintenance (using Trados’ MultiTerm software), translation memory (again for individuals and groups), access to CELEX (the full-text database of European Union legislation and directives), software for document comparison (to detect where changes have taken place), and also, of course, access to the Systran MT systems themselves.” (Hutchins 1999: 8)

Esta é uma estrutura sintática bastante ilustrativa que apresenta não só múltiplas referências, como também um inglês especializado, uma longa enumeração com exemplos que remetem para explicações em parêntesis e siglas. A quantidade de informação aqui presente dificultou uma proposta de tradução viável que conseguisse transportar as mesmas referências, mantendo o mesmo sentido. Isto porque uma longa construção sintática com múltiplos elementos em português pode, dependendo da sua construção, fazer com que o leitor se perca e não consiga reter a quantidade de conhecimento pretendida. Exemplos como o acima escolhido obrigaram a um repensar das opções tradutórias.

- “The latter are now available from English into Dutch, French, German, Greek, Italian, Portuguese, and Spanish; from French into Dutch, English, German,



Italian, and Spanish; from Spanish into English and French; and from German into English and French.” (Hutchins 1999: 8).

Este é outro exemplo de uma enumeração muito longa, com uma abundância de preposições para designar a direccionalidade da tradução. Foi necessário, mais uma vez, pensar qual a melhor forma de traduzir uma frase tão longa para que o leitor pressupusesse a mesma direccionalidade e não se perdesse na leitura.

#### **h) Características suprasegmentais**

A quantidade de características suprasegmentais presente no texto de Hutchins não é elevada e pode até, dada a organização visual do artigo, passar um pouco despercebida. No entanto, é importante referir que o autor faz algumas vezes uso do negrito e do itálico para dar ênfase às referências ou exemplos que dá.

Presente no seu texto está:

- i. “The first type of demand illustrates the use of MT for **dissemination**. It has been satisfied, to some extent, by machine translation systems ever since they were first developed in the 1960s.” (Hutchins 1999: 1)
- ii. “The second type of demand – the use of MT for **assimilation** – has been met in the past as, in effect, a by-product of systems designed originally for the dissemination application.” (Hutchins 1999: 2)
- iii. “Included in the latter are component programs to distinguish between homonyms (e.g. English words such as *light*, which can be a noun, and adjective or verb, and *solution*, which can be a mathematical or a chemical term) and to recognise the correct semantic relationships (e.g. in *The driver of the bus with a yellow coat*). (Hutchins 1999: 2-3)
- iv. “However, localisation is a multi-national and global industry, with its own organisation (Localization Industry Standards Association, based in Geneva) holding regular seminars and conferences in all continents (For details see *LISA Forum Newsletter*)” (Hutchins 1999: 8)

### **3.3. Crítica de Tradução**

**a) Coerência**

Como pode ser observado nos exemplos acima, a coerência do texto de partida pode ficar comprometida se as línguas de trabalho tiverem estruturas diferentes, como é o caso do português e do inglês. Para que o leitor não se perca na leitura e consiga seguir o raciocínio, foram adotadas medidas de adaptação das estruturas do inglês com estratégias de tradução em aceitabilidade em português. São exemplos:

- i. “As an alternative, the input text may be regularised (or ‘controlled’ in vocabulary and sentence structure) so that the MT system produces few errors which have to be corrected. Some MT systems have, however, been developed to deal with a very narrow range of text content and language style, and these may require little or no preparation or revision of texts” // “Como alternativa, o *input* textual pode ser regularizado (ou ‘controlado’ em termos de estrutura sintáctica e vocabulário) para que os sistemas de TA produzam poucos erros a necessitarem de correcção. No entanto, têm-se desenvolvido alguns sistemas de tradução automática para lidar com uma pequena variedade de conteúdos textuais e de estilo linguístico, o que requer pouca ou nenhuma preparação e revisão de texto.” (Hutchins 1999: 1)
- ii. “Systems have been developed by Smart for a number of major clients: Citicorp, Chase, Ford, General Electric, etc. Each incorporates a system for ‘normalising’ English documents.” // “A Smart tem desenvolvido sistemas para um número significativo de grandes clientes: a Citicorp, a Chase, Ford, a General Electric, entre outros. Cada um incorpora um sistema para ‘normalizar’ documentos em inglês.” (Hutchins 1999: 5)
- “Since the beginning of the 1990s, many other systems for personal computers have appeared.” // “Desde o início dos anos 90, aparecerem muitos outros sistemas para computadores pessoais. (Hutchins 1999: 9)

- “In particular, the Systran system, originally designed for translation only from Russian into English, is now available for a very large number of language pairs: English from and into most European languages (French, German, Italian, Spanish, Portuguese), Japanese, Korean, etc. Logos, originally marketed for German to English, is also now available for other languages: English into French, German, Italian and Spanish, and German into French and Italian. The Fujitsu ATLAS system, on the other hand, is still confined to translation between English and Japanese (in both directions).” // “Originalmente criado para a tradução exclusiva de russo para inglês, o sistema Systran está agora disponível para um vasto número de pares de línguas: de inglês para a maior parte das línguas europeias (francês, alemão, italiano, espanhol, português) e depois para japonês, coreano, entre outras. Também o sistema Logos que foi originalmente vendido de alemão para inglês, está agora disponível para outras línguas: de inglês para francês, alemão, italiano e espanhol, e ainda de alemão para francês e italiano. O sistema Fujitsu ATLAS, por outro lado, continua restringido à tradução entre inglês e japonês.” (Hutchins 1999: 9)

#### **b) Termos técnicos**

Como já foi referido, grande parte dos termos técnicos que a publicação de Hutchins apresenta não existe em português, pelo que foi difícil encontrar alternativas. Aquando da tradução foi necessário pesquisar quais os termos que tinha tradução e se já se encontravam consagrados, quais os empréstimos a utilizar e, caso fosse necessário, elaborar propostas de tradução pessoais para alguns termos. Dado que estamos perante um texto técnico, que fala concretamente sobre TA e tecnologias aplicadas à tradução, a má utilização ou tradução de muitos dos termos implicaria a incompreensão da leitura do texto de chegada e seu conteúdo. Pode dizer-se que existe um tom mais formal na proposta de tradução comparativamente ao texto de partida (frequente na utilização do presente do indicativo, por exemplo), o que ajuda à objetividade que o texto pede. Existe, como poderá ser observado, uma diferença de léxico terminológico que tenta ser colmatada havendo, no entanto, um grande número de empréstimos.

Entre os termos que trouxeram alguma dificuldade estão:

1. “Equally significant has been the use of MT for electronic mail and for ‘**chat rooms**.’” (Hutchins 1999: 11) // “Igualmente importante, tem sido o uso da TA para o *e-mail* e para as ‘**salas de chat**’.”

*Chat room*, ou *chatrooms* designam salas de conversação onde vários utilizadores, seja num *website*, aplicação ou programa podem conversar entre si, através da *internet*, ou outras aplicações num computador.

- Este é um termo cuja proposta de tradução é dificultada pela falta de adaptação em português. Existe a expressão “salas de conversação”, contudo, é expressão mais comum e coloquialmente utilizada em português-brasileiro, e não em português-europeu. Uma das razões pelas quais, provavelmente, o termo não é sequer utilizado em português é o facto de ser um tanto ou quanto extenso. Para além de que algo como “salas de conversação” não elucida o leitor sobre o conceito que dá origem ao termo. Uma sala remete-nos para um “compartimento principal de um edifício” ou “divisão principal de uma casa ou de um apartamento, onde geralmente se recebem visitas e se tomam as refeições” (Infopedia 2013). Já “conversação”, embora designe, exatamente, “acto de conversar” (Infopedia 2013), é um termo cuja utilização não é comum, e é preterida pelo termo “conversa” que designa “troca de palavras; diálogo” (Infopedia 2013). O que nos leva a sublinhar um aspeto muito importante: há que sempre ter em conta, entre muitos fatores, a probabilidade de acolhimento do termo proposto e a adequação ao vocabulário da especialidade em causa (Correia 2005: 7). Por esta razão, é uma expressão cujo empréstimo foi inevitável, contudo, não por completo. Em português é utilizada a expressão *chat*, consagrando o uso do empréstimo. *Chat* designa “ferramenta de comunicação, disponível *online*, utilizada para comunicar em tempo real” (Infopedia 2013). No entanto, embora este termo seja utilizado, tanto por escrito como oralmente, apenas a expressão *chat* não é suficiente para designar “rooms”, uma vez que a utilização do termo *chat*, sem mais nenhum nome a adjetivá-lo pode dar a entender que estamos perante uma única sala de chat, talvez de uso geral. Por esta razão, em português diz-se “salas de *chat*”, com a utilização de itálico para dar ênfase

ao empréstimo. Esta proposta de tradução foi convenientemente verificada junto de um especialista da área.

2. “Over the years this software company has constructed **controlled-language systems** for textile and insurance companies, mainly from Dutch to English (Van der Steen and Dijenborgh 1992).” (Hutchins 1999: 6) // “Ao longo dos anos esta empresa de *software* tem construído **sistemas de linguagens controladas** para seguradoras e empresas têxteis, principalmente de holandês para inglês (Van der Steen e Dijenborgh 1992).”

Este é um termo bastante específico, cuja significação é particularmente difícil de transmitir em português, uma vez que possui várias definições, com abordagens tanto para humanos, como para computadores. O termo “controlled language” é frequentemente abreviado para “CL” ou “CLs” (“Controlled languages”, para designar o plural). A sua conceção adveio da dificuldade em compreender e traduzir a língua natural, devido a problemas de ambiguidade e complexidade (Somers 2003: 245). A Linguagem Controlada (ou “linguagens controladas”), como o nome indica, controla e restringe a língua que é utilizada em textos de autor. Ao limitar o léxico, a gramática e o estilo, é proporcionada uma leitura e compreensão mais eficaz e menos confusa. E, acima de tudo, “(...) enhances the efficiency and accuracy of the tasks associated with technical documentation, and improves the quality of human- and machine-translated text” (Somers 2003: 245). No entanto, é importante referir que a sua definição depende da sua utilização, e de quem a solicita (daí que seja importante sublinhar que existem inúmeras definições para uma “linguagem controlada”). Tanto pode ser incorporada num sistema, como ser proposta por grandes empresas para diferentes tipos de documentação, como ser utilizado em conjunto com um *software* de verificação de novos textos, entre outros – é isto que definirá CL em questão. Apesar de haver, portanto, diferentes definições de “Linguagem Controlada”, existem dois objetivos comuns a qualquer uma: a redução do número de palavras (para que exista uma correspondência literal entre a palavra e o seu conceito, excluindo casos em que o conceito pode ser expresso por outras palavras ou quando uma palavra pode exprimir mais do que um conceito) e eliminar a possibilidade de ambiguidade em construções

sintáticas complexas. Estas diretrizes podem ser caracterizadas como “human-oriented” ou “machine-oriented”, ou seja, podem ser concebidas e utilizadas para facilitar a compreensão de um texto por humanos ou por computadores – reforçando qual os pontos fortes da CL. No entanto, é ainda importante mencionar que a existência desta restrição veio melhorar a consistência e a reutilização do texto de partida (Somers 2003: 248), que juntamente com ferramentas de tradução se torna ainda mais eficaz.

- Esta é uma expressão complexa, uma vez que o nome em inglês suscita algumas dúvidas: “language” tanto pode referir-se a língua como linguagem. E mesmo que a proposta de tradução obedeça ao conceito (e ao contexto) que a principia, continua a ser um termo que pode deixar o leitor sem perceber muito bem a que se refere, isto porque é um termo técnico cuja compreensão só é conseguida, caso o leitor se debruce a fundo em leituras de textos paralelos. Depois de uma pesquisa inicial, não foi difícil perceber que o termo deriva do conceito linguístico “língua natural” – termo que designa um sistema de “comunicação verbal que se desenvolve espontaneamente no interior de uma comunidade. Este termo opõe-se a língua artificial.” (Iltec – Instituto de Linguística Teórica e Computacional 2013). Em português existe a distinção, importante, entre língua e linguagem, e foi esta diferença que foi decisiva aquando da tradução. Na língua portuguesa, “língua” designa não só “o órgão mais ou menos musculoso e móvel, tipicamente alongado, existente na cavidade bucal, que serve para a degustação e a deglutição e, no homem e em alguns animais, tem um papel importante na articulação dos sons” (Infopedia 2013), como a “capacidade humana de comunicar através do uso sistemático e convencional de sons, sinais ou símbolos escritos. O termo é utilizado para exprimir outros conceitos como os meios de comunicação dos animais ou os sistemas de programação em informática” (Iltec – Instituto de Linguística Teórica e Computacional 2013). Contudo, “linguagem” significa “qualquer sistema ou conjunto de sinais convencionais, fonéticos ou visuais, que servem para a expressão dos pensamentos e sentimentos; modo particular pelo qual uma pessoa se exprime; oralmente ou por escrito; maneira de falar; estilo; sistema de representação que os membros de uma comunidade linguística usam como principal meio de comunicação, falado ou escrito; língua” (Infopedia 2013). Se o termo “controlled language” se refere à necessidade de adotar estratégias de controlo para textos que serão traduzidos por sistemas de TA, de

forma a resolver os problemas que surgem (como os de ambiguidade, inferência, entre outros), presume-se que estes problemas surgem de uma “língua natural”. Aqui entra então a opção “linguagem” e não “língua” dado que, como observámos, linguagem exprime exatamente o sistema complexo que nós, humanos, utilizamos como meio de comunicação, a forma pela qual nos exprimimos por escrito e oralmente. Pretende-se, pois, controlar esta linguagem, daí que o termo proposto seja “linguagens controladas”, e não “línguas controladas”. A opção “línguas controladas” podia levar o leitor a pensar que se trata de uma língua em específico, quando na verdade é todo o conjunto de sinais pelos quais uma pessoa se expressa, independentemente da língua materna que fala. Por esta razão, a opção mais clara e apropriada será “linguagens controladas”. Porém, é importante salientar que para a escolha final, contribuiu o facto de ser uma expressão cuja proposta de tradução não é completamente estranha. Exemplo disso é a descrição de Elisabete Ranchhod que, sobre a linguagem controlada escreve “(...) podemos referir a existência de sistemas (...) que, através do uso de analisadores de uma linguagem simplificada (*controlled language*), forçam uma consistência no estilo da linguagem (...) simplificando significativamente a tradução automática subsequente.” (Ranchhod 2001: 238).

3. “The final section compares the types of translations that can be met most appropriately by human and by machine (**and computer-aided**) translation respectively.” (Hutchins 1999: 1) // “A secção final compara os tipos de tradução que podem ser apresentados de forma mais adequada, tanto por uma tradução humana, como por uma tradução automática (e ainda por uma **tradução assistida por computador**).”

Começando por uma das palavras-chave da publicação de Hutchins, “computer-aided translation” é um dos termos mais frequentes e indissociáveis do campo da TA. Como o nome indica é um tipo de tradução que é processada, gerida e “auxiliada” através de ferramentas computacionais. Este termo é frequentemente abreviado para o conhecido acrónimo CAT – computer-aided translation – e aparece em grande parte da bibliografia de TA associado a ferramentas de tradução, nomeadamente como “CAT tools”. Isto porque, sendo um tipo de tradução que não é completamente automático,

confere algum controlo ao tradutor, “assistindo” a tradução através destas ferramentas – “CAT tools for translation and software localization nowadays offer, almost without exception, a bundle of functions (...)” (Munday 2009: 117). Dentro do leque de ferramentas CAT, destacam-se os importantíssimos *softwares* de localização e de trabalho do tradutor.

▪ Este é dos poucos termos que podemos encontrar, que não só já têm uma tradução para português, como possuem também um uso legitimado em português. Tanto é que o acrónimo pelo qual, tanto a tradução como as ferramentas são conhecidas – CAT – tem uma correspondência em português. “Computer aided-translation” é já uma área de estudo com destaque dentro da área de tradução, especialmente em Portugal. Como exemplo disso temos já a integração do estudo deste tipo de tradução como disciplina na Faculdade de Letras da Universidade de Lisboa<sup>15</sup>. Por esta razão, não só o termo traduzido é de conhecimento comum, como é passível de uso o seu acrónimo em português – TAC – “tradução assistida por computador”. Prova da consagração do uso é também a entrada nas novas edições de dicionários como a Porto Editora, que no seu dicionário de Inglês-Português já possui uma entrada para “computer-aided” que é exatamente “assistido por computador (Porto Editora 2009: 177). Como foi referido acima, este é um tipo de tradução que não é completamente automático. Isto porque a tradução é efetuada através do auxílio de algumas ferramentas, conferindo ao tradutor autonomia para melhorar, facilitar e agilizar o processo de tradução. A tradução pela qual é sobejamente conhecida (TAC) é bastante clara para o leitor, pelo que não é necessário acrescentar nada mais além da breve descrição que figura acima.

4. “This survey of the present demand and use of **computer-based translation** software concentrates on systems designed for the production of translations of publishable quality, including developments in controlled language systems, translator workstations, and localisation;” (Hutchins 1999: 1) // “O presente

---

<sup>15</sup> Seminário lecionado pelo Professor André Eliseu, na Faculdade de Letras da Universidade de Lisboa. Programa pode ser consultado em: <http://www.fl.ul.pt/dlgr-docentes/913-dlgr-programas-traducao-assistida-por-computador> (2013)



estudo sobre a procura e uso do *software* **de tradução baseado em computador** [computer-based] concentra-se nos sistemas criados para a produção de traduções de qualidade publicável, incluindo desenvolvimentos em sistemas de linguagens controladas, *software* de trabalho do tradutor e de localização.” – “Whereas, with MT systems translators see themselves as subordinate to the machine, in so far as they edit, correct or re-translate the output from a computer, with translation workstations (or workbenches) the translators are in control of **computer-based facilities.**” (Hutchins 1999: 2) // “Este ‘*software* de trabalho de tradução’ é mais atrativo para tradutores humanos, enquanto os tradutores dos sistemas de TA se vêem subordinados à máquina no que diz respeito à edição, correção e à re-tradução do *output* de um computador. Com o *software* de trabalho de tradução, os tradutores têm controlo absoluto sobre os **recursos baseados em computador [computer-based].**”

- Esta é uma opção tradutória bastante arriscada, uma vez que existe pouca bibliografia, tanto em português como inglês, em que o termo apareça referenciado. É apenas nos textos de Hutchins que encontramos este termo. Existem duas ocorrências para este termo: “computer-based translation *software*” e “computer-based facilities”. Sendo que “facilities” designa “funções, recursos; instalações, serviços (Porto Editora 2009: 311) e numa tradução à letra o termo parece indicar que tanto os serviços, como o *software* partilham características que podemos encontrar num computador, a tradução literal não parece uma má opção. Vejamos o contexto. Na primeira ocorrência, que aparece no resumo da publicação, diz-nos o termo que o *software* de tradução “computer-based” concentra-se nos sistemas que são criados para a produção de traduções de uma qualidade que se pretende publicável. Presumimos que o *software* de tradução é idêntico ao de computador que, para ser “computer-based” deve ser semelhante a um *software* operacional, i.e., um tipo de software que integra e trabalha diretamente com o sistema operativo – o computador. O mesmo acontece na segunda ocorrência, que diz que os serviços ou operações são “computer-based”, o que parece indicar que determinado sistema ou aplicação possui as mesmas funcionalidades que um computador. Por esta razão, opta-se pela proposta de tradução literal, à falta de

bibliografia que suporte outra escolha, e sendo que é uma opção que, apesar de condicionada, faz sentido no contexto que possuo.

5. “By the 1980s, however, advances in **computational linguistics** allowed much more sophisticated approaches, and a number of systems adopted an indirect approach to the task of translation.” (Hutchins 1999: 2) // “Por volta dos anos 80 do século XX, contudo, deu-se um avanço na **linguística computacional** que permitiu abordagens muito mais sofisticadas, e vários sistemas adotaram uma abordagem mais indireta à tarefa da tradução.
- A linguística computacional é um ramo da linguística que fala das línguas naturais, tanto escritas como faladas. É uma área interdisciplinar que surge frequentemente como sinónimo de NLP (*Natural Language Processing*) (Mitkov 2005: 747, i.e., o Processamento de Linguagem Natural - PLN. O processamento é feito através de dispositivos automáticos, obrigando a “uma descrição sistemática e completa das línguas a tratar, de modo a evitar ou, pelo menos, reduzir as falhas de processamento devidas à insuficiência dos dados linguísticos” (Ranchhod 2001: 13). O léxico surge, portanto, como o motor deste processo, uma vez que é sobre os recursos lexicais que recai a importância do tratamento automático na manipulação da linguagem humana. Não é um termo que possa suscitar problemas, no entanto, convém esclarecer o leitor sobre a área e o que esta incorpora. Em Portugal, nomeadamente em Lisboa, podemos encontrar o ILTEC – o Instituto de Linguística Teórica e Computacional, a Faculdade de Ciências da Universidade de Lisboa e o INESC – Instituto de Engenharia de Sistemas e Computadores como referências importantes – historicamente.
6. “Based on experience over many years in developing knowledge-based MT and experimenting with speech translation and **corpus-based methods**, members of the group at Carnegie-Mellon University have developed an architecture for the rapid production of usable MT systems for specific clients in some less common languages, such as SerboCroat and Haitian Creole (Frederking et al. 1997).” (Hutchins 1999: 6) // “Baseado numa experiência efetuada ao longo dos anos,

esta consiste em desenvolver TA baseada em conhecimento, e que faz experiências com a tradução de discurso e **métodos baseados em corpora** [**corpus-based methods**]. Membros do grupo na universidade de Carnegie-Hall têm desenvolvido uma estrutura para a produção rápida de sistemas de TA usáveis para clientes específicos, em línguas pouco comuns, como o servo-croata ou o crioulo do Haiti (Frederking et al. 1997).”

Utilizado com frequência, este é um termo geralmente associado não só a um tipo de TA, mas também a desenvolvimentos dentro da área que vieram melhorar a qualidade de ferramentas de auxílio ao tradutor (especialmente dicionários bilingues e multilingues). Diz-se que *corpus* é uma coleção de textos “stored on a computer, sometimes analysed automatically or semi-automatically” (Somers 2003:105). Esta coleção de textos consiste em conversas do dia-a-dia, emissões de rádio, livros já publicados, entre outros elementos, que servem para analisar de forma refletida uma língua. Este tipo de tarefa é feito através de uma estrutura semelhante à memória de tradução, contudo, a sua particularidade consiste no facto de alinhar de forma paralela o *corpus* (em inglês este recurso é frequentemente designado como “aligned parallel corpus”). É uma ferramenta que facilita a compreensão do texto, permitindo ao mesmo tempo que qualquer informação linguística (quer do texto de partida, quer do de chegada) seja utilizada na comparação de outras traduções para o texto original.

- Embora seja um termo técnico, associado ao campo da linguística computacional, nomeadamente ao Processamento de Linguagem Natural, este não foi um termo difícil de traduzir. Uma das razões principais é o facto do nome “corpus” ser sobejamente conhecido e ter, inclusive, uma utilização bastante comum. *Corpus* designa a “compilação de documentos ou informações relativos a uma disciplina ou tema; conjunto finito de enunciados representativos de uma determinada estrutura” (Infopedia 2013). Dado que ambas as definições, especialmente a última, encaixam perfeitamente na designação que o termo adquiriu na TA, e dado de que se trata de um termo de utilização universal, optou-se por “métodos baseados em *corpora*”. Dado o contexto e as informações encontradas sobre o termo é possível perceber que estamos perante

métodos de pesquisa “baseados” em *corpora*, que são suscetíveis de serem analisados automaticamente – dando-se ênfase à preposição “em”. Um *corpus*, enquanto um conjunto de fontes relativas a um domínio (Iltec – Instituto de Linguística Teórica e Computacional 2013), refere-se apenas a uma só disciplina ou tema. Razão pela qual a opção escolhida foi o plural de *corpus*, i.e. *corpora*, para ficar claro que são métodos que podem ser aplicados a mais de um objeto de estudo. É ainda importante salientar que o termo “TA baseada em *corpus*” já tem uma tradução, que embora ainda não seja de utilização comum, foi encontrada numa tese orientada pela Doutora Palmira Marrafa, *Para uma Tradução Automática Baseada em Conhecimento: Especificação da Modificação e da Predicação Adjectiva*, da aluna Náheda Ibrahim (Ibrahim 2010: 15).

7. “There is growing interest in such multilingual applications worldwide. The application that has received most attention has been ‘**cross-language information retrieval**’, i.e. software enabling users to search foreign language databases in their own languages.” (Hutchins 1999: 12) // “Existe um interesse crescente neste tipo de aplicações multilingues em todo o mundo. A aplicação que tem recebido mais atenção, tem sido a **extração de informação multilingue [cross-language]**, i.e. um *software* que permite aos utilizadores pesquisarem bases de dados de línguas estrangeiras, na sua própria língua.”

Frequentemente conhecido como CLIR – Cross-Language Information Retrieval, veio revolucionar a área de tradução automática ao aliar dois campos bastante importantes e complexos: o campo da recuperação de informação (que se insere na área de informática, uma vez que diz respeito ao armazenamento de documentos) e o da TA. Desenvolvido em 1970 por Salton, CLIR designa a recuperação ou reaquisição de informação numa língua que não é a língua-mãe do utilizador, mas que este consegue encontrar através de uma dúvida sua expressa na sua língua. Citando Mitkov, “With information retrieval, cross-language information retrieval researches robust language abstraction techniques which are domain independent. It tries to map the vocabular used by the documents and the vocabular used in the query to a space in which the computer can recognize similarity.” (Mitkov 2005: 702). A capacidade de procurar uma

informação que se encontra numa língua, através de questões na nossa própria língua é possível através da manutenção e integração de dicionários bilingues e multilingues.

- Este foi um dos exemplos que, por ser bastante técnico e não possuir uma tradução já dicionarizada, teve de ter uma proposta de tradução devidamente fundamentada. Dado que estamos perante o campo de extração de informação, a expressão “information retrieval” não apresentou qualquer dúvida, uma vez que significa exatamente “recuperação; reaqisição; extração de informação” (Porto Editora 2009: 913). Contudo, o grande desafio é apresentar uma proposta de tradução para “cross-language”. “Cross-language” numa tradução à letra, “línguas cruzadas”, no entanto, no léxico português é uma expressão que não existe, e que não tem nenhum significado. Foi proposto então o termo “multilingue” que designa algo que “possui muitas línguas; que fala muitas línguas” (Infopedia 2013). Como estamos perante uma forma de aceder à informação pretendida que se encontra numa língua específica, através de outra diferente (Mitkov 2005: 700) estamos perante uma aplicação que contém vários tipos de informação num grande número de línguas. Por esta razão, em vez de optar por algo como “extração de informação de várias línguas” que era demasiado comum para corresponder a um termo tão técnico, e não podendo optar apenas por “extração de informação”, uma vez que estaria assim apenas a referir-me à área da informática sem a sua interligação à TA, optou-se por “recuperação de informação multilingue”.

8. “However, possibly the best known success story for **custom-built MT** is the PaTrans system developed for LingTech A/S to translate English patents into Danish. The system is based on methods and experience gained from the Eurotra project of the European Commission (Ørsnes et al. 1996).” (Hutchins 1999: 6) // “No entanto, a história de sucesso mais conhecida de um **sistema de TA ‘feito à medida’ [custom-built]** é possivelmente o sistema PaTrans desenvolvido pela LingTech A/S para traduzir patentes de inglês para dinamarquês. Este sistema é baseado em métodos e experiências no âmbito do projeto Eurotra da Comissão Europeia (Ørsnes et al. 1996).”

- Tal como aconteceu com o termo técnico “computer-based”, este é outro exemplo de um termo que não tem qualquer referência bibliográfica além da publicação de Hutchins. “Custom-built” designa exatamente algo que “é feito à medida” (Porto Editora 2013: 211), e tem um uso mais coloquial que escrito. Tal como aconteceu anteriormente, à falta de textos paralelos e suporte bibliográfico, temos que analisar bem o contexto para perceber em que medida uma tradução literal faz, ou não, sentido. No texto do autor, “custom-built” é um termo que qualifica um sistema com as características que a empresa pediu. Isto significa que este sistema foi feito consoante o pedido de um cliente em particular, com as características que pretendia. *Custom-built* é associado quase sempre a um pedido específico de determinada empresa ou cliente e, por esta razão a opção tradutória não oferece outra alternativa. Foi também consultado o motor de busca Google, no entanto, não foi encontrado nenhum *site* creditado que possa ser referido como fonte.

9. “For many years, the systems were based primarily on **direct translations** via bilingual dictionaries, with relatively little detailed analysis of syntactic structures. By the 1980s, however, advances in computational linguistics allowed much more sophisticated approaches, and a number of systems adopted an **indirect approach** to the task of translation. (...) In one the abstract representation is designed to be a kind of **language-independent ‘interlingua’**, which can potentially serve as an intermediary between a large number of natural languages.” (Hutchins 1999: 2-3) // “Por muitos anos, os sistemas foram inicialmente baseados em **traduções diretas** via dicionários bilingues, com relativamente pouca análise detalhada das estruturas sintáticas. Por volta dos anos 80 do século XX, contudo, deu-se um avanço na linguística computacional que permitiu abordagens muito mais sofisticadas, e vários sistemas adotaram uma abordagem mais **indireta** à tarefa da tradução. (...) Numa, a representação abstrata pretende ser uma espécie de **‘interlíngua’ independente**, que pode potencialmente servir como um intermediário entre um grande número de línguas naturais.”

- Não sendo necessário explicar em que consiste a abordagem direta, uma abordagem por interlíngua requer uma explicação um pouco mais sólida. Utilizando igualmente técnicas de inteligência artificial, este tipo de sistema tem como objetivo fazer as representações do significado do texto da língua de partida, independentemente da língua de chegada. A abordagem por transferência é considerada uma das arquiteturas clássicas da TA (Somers 2003: 122), composta por três etapas (a análise, transferência e geração) e cada uma pretende eliminar os problemas mais sistemáticos da TA, como a ambiguidade por exemplo. Todas estas opções tradutórias são complementadas pela preposição “por”, uma vez que esta designa exatamente o modo de tradução adotado, e por conseguinte, quais os resultados que daí advirão.

10. “In late 1997 it launched its multilingual service for the translation of **electronic mail**, Web pages and attached files.” (Hutchins 1999: 10) // “No final do ano de 1997, foram ainda lançados os serviços multilingues para a tradução de *e-mails*, páginas *Web* e anexos.”

“Electronic mail” ou correio eletrónico é a forma de transmitir e enviar mensagens, normalmente através de redes de comunicação. As mensagens são digitadas, normalmente, através de um teclado de computador e na mensagem pode seguir em anexo qualquer tipo de ficheiro, desde imagens a vídeos. Citando Neves “Todos os sistemas de acesso à *internet* dispõem de correio eletrónico, permitindo a troca de mensagens com utilizadores de qualquer sistema” (Neves 2009: 104).

- Independentemente da opção tradutória, este é um termo que não deixa qualquer margem de dúvidas ao leitor, uma vez que este consegue perceber que se trata de correspondência eletrónica. Aqui, a dificuldade de escolha recai em qual será o termo apropriado, e não na procura de um termo em português que seja uma opção viável. Em português existem várias opções: “correio eletrónico”, *e-mail* ou ainda *mail*, sendo as últimas as abreviações de “electronic mail”. No entanto, coloquialmente não é frequente alguém dizer que recebeu “um correio eletrónico” da pessoa X ou Y ou perguntar “qual

é o seu correio eletrónico?”. É por esta razão que a língua portuguesa se apropriou, também, do termo *e-mail* (ou até *mail* coloquialmente utilizado em linguagem de jovens). *E-mail* tem um uso quicá muito mais frequente e vulgar. É por esta razão que, à semelhança da tradução literal “correio eletrónico”, também podemos encontrar na Porto Editora Online uma entrada para *e-mail* que, além de possuir a definição acima escrita, também designa: “mensagem enviada através deste sistema; local, no computador, onde chegam e são mantidas as mensagens que alguém recebe por este sistema; caixa de correio eletrónico; expressão que identifica um utilizador numa rede de computadores, permitindo o envio e a receção de mensagens por este sistema; endereço eletrónico” (Infopedia 2013). Ambas as opções estão, por conseguinte, corretas e qualquer uma pode ser aceite. Porém, optei por *e-mail*. Não só a sua definição é muito mais completa, como em português é a expressão atualmente mais utilizada e frequente. Esta proposta justifica-se por ser um empréstimo consagrado, que inclui um pseudoprefixo, que resulta da truncação de um termo mais longo (Correia 2005: 6) que provem de “electronic”, e que se li “i”.

11. “The fourth type of MT application – as **components of information access systems** – is the integration of translation software into: (...)” (Hutchins 1999: 2) // “O quarto tipo de aplicação de TA – como **componentes de sistemas de acesso à informação** – é a integração de *software* de tradução em: (...)”

- À semelhança do termo “TA para intercâmbio”, nenhuma outra publicação exibiu este termo. Por esta razão, e estudando o contexto onde se inseria a expressão e conduzindo uma pesquisa na *internet*, concluímos que estamos perante uma espécie de “sistema de informação”, que consiste num sistema que permite gerir e manipular dados, recorrendo ou não a ferramentas de tecnologia de informação. Dado que permitem armazenar uma grande quantidade de informação importante, são componentes frequentemente utilizadas para sistemas informáticos de áreas como a Saúde, o Direito, Gestão, entre outros. Como suporte bibliográfico encontramos, por exemplo, uma ata sobre *Os sistemas de Gestão da Informação nos Hospitais Públicos Portugueses* (BAD – Associação Portuguesa de Bibliotecários, Arquivistas e



Documentalistas 2013) que fala exatamente da importância destes sistemas para a preservação e manutenção de todo o tipo de informações.

12. “The system, at present running live for English into Spanish, demanded techniques otherwise found mainly in Internet applications” (Hutchins 1999: 7) // “O sistema, de momento a funcionar em direto de inglês para espanhol, exigia técnicas que eram encontradas principalmente em muitas **aplicações para a internet**.”

Não é uma opção tradutória que possa suscitar dúvidas. Contudo, em muitos dos termos do texto de Hutchins, foi importante procurar o seu significado, uma vez que nem todos dizem respeito à área da tradução, mas sim à área da Informática. Aplicações são vários tipos de serviços que não podem funcionar sem o sistema operativo que os rege. Citando Jorge Neves, aplicações podem ser “processadores de texto, bases de dados ou folhas de cálculo. De certa forma ocupam a posição de topo do *software*, porque não podem funcionar sem o sistema operativo e os programas utilitários do sistema” (Neves 2009: 26).

- Estamos perante uma expressão que designa programas especialmente concebidos para determinado sistema, e com uma utilização específica em mente. Segundo uma fonte especialista na área de informática, em português a preposição “para” indica, neste caso, o lugar de destino da aplicação. E como é perceptível, estamos perante aplicações (ou programas) que foram concebidos especialmente para a *internet*. A nível coloquial utiliza-se a expressão “aplicações da Internet”, no entanto, existe uma grande diferença entre dizer “aplicações **de internet**” e “**para a internet**”. Caso figurasse o primeiro termo, deixaríamos de referir todas as aplicações concebidas, de propósito, para a *internet*, e estaríamos a designar aplicações que pertencem à *internet*.

13. “One of the fastest growing areas for the use of computers in translation is in the industry of software localisation.” (Hutchins 1999: 8) // “Uma das áreas de

maior crescimento relativas ao uso de computadores para tradução é a indústria do ***software de localização***.”

À semelhança do *software* de trabalho do tradutor e da memória de tradução, o *software* de localização é outra das importantes componentes das ferramentas CAT, em português TAC – Tradução Assistida por Computador. Muitas das suas características têm um funcionamento semelhante à memória de tradução e ao *software* de trabalho do tradutor. O nome deste *software* deriva da expressão *locale*, que é uma palavra que possui diferentes possibilidades de significado. Geralmente designa “all characteristics of the combination of a language and a region or country.” (Somers 2003: 67). Isto significa que o objetivo primário desta ferramenta é adaptar manuais para produtos, os próprios produtos e o *software* para um público “local”. A própria associação de Localização (a LISA – *Localisation Industry Standards Association*) designa esta ferramenta como algo que “involves taking a product and making it linguistically, technically, and culturally appropriate to the target locale where it will be used and sold.” (Localisation Industry Standards Association *apud* Somers 2003: 67). Adaptar um produto linguisticamente a determinado mercado, com todas as tarefas que são adjacentes, não é a única particularidade desta ferramenta. Para além deste conceito, o *software* de localização tem outras particularidades - “(...) the program code by extracting the translatable text (...) for translation in a safe environment and eventually reinsert the translated version in the right places in the right files” (Munday 2009: 120). À semelhança da memória de tradução, que memoriza o termo necessário e permite que seja, mais tarde, consultado, também o *software* de localização possui essa funcionalidade. Para além disto, uma das ferramentas principais de que dispõe consiste em simular o *interface* traduzido – aplicação através da qual o utilizador comunica com o computador (Neves 2009: 26) – e “localizar” qualquer caso digno de registo.

- Uma vez que estamos perante um *software* cuja função principal consiste em adaptar as suas ferramentas a um público “local”, e as suas funcionalidades são idênticas às de memória de tradução a qual, em vez de memorizar, se especializa em localizar os termos pesquisados, não me pareceu que fosse uma proposta de tradução

incoerente optar por uma tradução literal. A sua funcionalidade é explicitada em muita bibliografia. O único risco de incongruência que pode ocorrer é o facto de o leitor poder concluir que se trata de um *software* que localiza algo ou alguém. Contudo, propostas como “*software* de local” ou “*software* para um local”, não seriam opções tradutórias possíveis, uma vez que este programa não se limita apenas a adequar as suas ferramentas a um tipo de público específico, local. Este *software* também possui e permite as mesmas funcionalidades que a memória de tradução, pelo que as opções descritas seriam incongruentes e não elucidaria o leitor sobre o que, afinal, consiste.

14. “These advantages have been recognised by other multinational companies, and the use of ‘controlled languages’ is increasing: for example, the Caterpillar Corporation has devised its own form of English to facilitate translation in a **knowledge-based MT system** being developed for it at the Carnegie-Mellon University (Mitamura and Nyberg 1995).” (Hutchins 1999: 5) // “Estas vantagens têm sido reconhecidas por outras empresas multinacionais, e o uso de ‘linguagens controladas’ está a aumentar: por exemplo, a Caterpillar Corporation [Corporação Caterpillar] criou a sua própria forma de inglês para facilitar a tradução **num sistema de TA baseado em conhecimento**, a ser desenvolvido pela Universidade Carnegie-Mellon (Mitamura e Nyberg 1995).”

Também conhecido como “interlingua-based MT”, o sistema de tradução “knowledge-based” funciona à base de uma representação do sentido do texto (abreviado na sigla TMR – *text-meaning representation*) que indica o resultado de análise de um qualquer *input* de texto em qualquer língua que o sistema, leia-se multilingue, suporte. Mas o que caracteriza este conceito, ao ponto de lhe conferir este nome peculiar, é o facto de “the meaning of the input text (...) is represented in the TMR as elements which must be interpreted in terms of an independently motivated model of the world (or *ontology*).” (Mitkov 2005: 520). Ontologia é o modelo que determinado falante tem acerca do mundo, sendo que este modelo é constituído por conceitos ou mapeamentos que, consoante o indivíduo, são organizados e hierarquizados de forma diferente e única (Mitkov 2005: 520). Neste tipo de sistemas o saber é, portanto, definido através de

mapeamentos dos termos interpretados dentro do conceito da ontologia – que permite a categorização de uma palavra, consoante o seu significado e as propriedades que manifesta. Esta particularidade é uma forma de reduzir a ambiguidade do texto, à semelhança dos sistemas de linguagens controladas, uma vez que delimita o sistema a determinadas áreas.

- Esta tradução é uma proposta pessoal. Como foi referido acima, um sistema baseado em conhecimento, utiliza o conhecimento que deriva da informação lexical, sintática, semântica e pragmática para a organizar e classificar. É uma forma de organizar o saber por áreas, delimitando as opções tradutórias. Por esta razão, é um dos poucos exemplos em que uma tradução literal, embora pareça estranha, faz sentido. Não só corresponde à ideia central do sistema, como é uma equivalência direta a nível do significado que o seu nome possui. É uma opção um tanto ou quanto estranha, uma vez que o utilizador não é elucidado sobre o tipo de conhecimento em que o sistema se baseia, no entanto, é esse o seu conceito principal. Aquando da justificação desta escolha foi encontrado o mesmo termo na já referida tese de Náheda Ibrahim, *Para uma Tradução Automática baseada em Conhecimento: especificação da modificação e da predicação adjectiva*. (Ibrahim 2010: 43).

**15. “With the third type – MT for interchange – the situation is changing quickly.**

The demand for translations of electronic texts on the Internet, such as Web pages, electronic mail and even electronic ‘chat’ lists, is developing rapidly.” (Hutchins 1999: 2) // “Com o terceiro tipo – o **de TA para intercâmbio** – a situação está a mudar rapidamente. A procura de traduções de textos eletrónicos na *internet*, como de páginas *Web*, *e-mail* e mesmo ainda de listas eletrónicas de ‘chat’, está a desenvolver-se rapidamente.

- Sobre este termo, infelizmente, não foi possível encontrar nenhum texto que auxiliasse a opção tradutória, senão a publicação de Hutchins. Sem qualquer texto paralelo que nos elucidasse melhor sobre os conceitos e objectivos deste tipo de procura

de TA, limitamo-nos à análise a partir do contexto, tal como aconteceu com outros termos. Podemos depreender, da descrição de Hutchins, que este tipo de procura é o mais atual, uma vez que se apoia mais sobre as tecnologias mais recentes. Sabemos que estamos perante a tradução de meios de comunicação de natureza tecnológica, como salas de *chat*, troca de *e-mails*, páginas *Web*, entre outros. É um tipo de tradução que se quer rápida e “na hora”, daí que o tradutor humano raramente seja equacionado, uma vez que é o computador que pode funcionar com rapidez *online*. Adquiriu este nome, provavelmente, porque as traduções pedidas dizem respeito a tipos de comunicação de um para um, via estes meios de comunicação que permitem este tipo de “troca”. Sendo que o termo “troca” é deveras coloquial para cunhar um tipo de procura de TA que, segundo o autor, se tem tornando bastante emergente e tem tido alguma saída, optou-se pelo termo “intercâmbio” – mais formal e que exprime, exatamente “trocar, permutar, inverter” (Infopedia 2013).

16. “Translation workstations combine **multilingual word processing**, means of receiving and sending electronic documents, OCR facilities, terminology management software, facilities for concordancing, and in particular ‘translation memories.’” (Hutchins 1999: 7) // “O *software* de trabalho de tradução combina o **processamento de palavras multilingue**, com meios de receber e enviar documentos eletrónicos, serviços OCR [Optical Character Recognition], *software* de controlo de terminologia, funcionalidades que permitem fazer concordâncias, e em particular a ‘memória de tradução’.

- Um processador de palavras é uma aplicação de computador que permite a elaboração de um documento e tudo o que esta tarefa implica, i.e., a revisão, formatação e até a impressão de documentos. Entre estas funcionalidades, para a elaboração de documentos, o utilizador dispõe ainda de um corretor ortográfico. Ferramentas de que qualquer pessoa dispõe e utiliza frequentemente numa plataforma como o Microsoft Word®, por exemplo. De acordo com a Infopedia, é um “programa de computador que é usado para criar, modificar, guardar e imprimir texto” (Infopedia 2013). É um dos poucos termos que tem uma tradução, embora não tenha uma utilização vulgar como

*software* ou *hardware*. E como muitos programas como o Word são utilizados diariamente, e o utilizador executa essas tarefas de forma tão mecânica, dificilmente assimila o facto de haver uma aplicação que permite todas estas funcionalidades. Como possui uma tradução reconhecida na área de informática, não foi uma opção tradutória difícil, no entanto, existe a necessidade de alertar o leitor para o respetivo conteúdo.

17. “At the same time, many MT vendors have been providing **network-based translation services** for on-demand translation, with human revision as optional extras.” (Hutchins 1999: 10) // “Ao mesmo tempo, muitos vendedores de TA têm fornecido **serviços de tradução em rede [network-based]** para traduções encomendadas, com revisão humana como extras opcionais.”

- À semelhança de termos como “custom-built” ou “computer-based”, este é um termo que também não foi possível encontrar em nenhuma da bibliografia selecionada. Como tal, pensa-se que o termo foi cunhado por Hutchins, para indicar que os serviços de tradução em questão foram baseados em “rede” – i.e., através de “uma rede de computadores que estão ligados entre si” (Infopedia 2013), do inglês “network”. No entanto, ao contrário dos casos referidos, optei por não incluir a tradução literal do termo “based”, uma vez que “serviços de tradução baseados em rede”, é uma expressão estranha que não elucida de todo o leitor. Se “network” já exprime algo que está ligado “em rede” ou um trabalho em rede, a utilização da expressão “baseada” torna-se desnecessária e obtusa, podendo inclusive chegar a pensar-se que serviços “baseados em rede” indicam algo mais do que a expressão quer, na verdade, dizer.

18. “It was only in the 1980s that developments in **speech recognition** and synthesis made spoken language translation a feasible objective.” (Hutchins 1999: 12) // “Foi só nos anos de 80 do século XX que desenvolvimentos no **reconhecimento de fala** e síntese fizeram da tradução de língua falada, um objetivo exequível.”

- Espelho de uma das tecnologias mais importantes dentro da área da TA, “speech recognition” designa, em português, o “reconhecimento de fala”. Os sistemas de reconhecimento de fala são aqueles que “permitem a atribuição de estruturas linguísticas a sequências acústicas de fala” (Ranchhod 2001: 195). É importante reconhecer qual o seu objetivo primário, uma vez que é isto que distingue os vários tipos de sistemas, e facilmente se confunde algo como sistemas de reconhecimento de fala, com sistemas de comandos vocais. Ambos implicam o aparelho vocal, contudo, enquanto o sistema de comandos de voz implica apenas algumas dezenas de entradas por ser mais limitado a nível de vocabulário, o sistema de reconhecimento de fala relaciona as unidades acústicas com uma estrutura dotada de sentido, exigindo milhares de entradas. Citando Fernando Martins “(...) consideramos que a diferença entre os dois conceitos reside no facto de os sistemas de compreensão de fala acederem a informações de ordem semântica (Martins *apud* Ranchhod 2001: 196). Uma das etapas de um sistema de reconhecimento de fala consiste na produção de uma representação através de um gerador de frases. Esta representação é convertida pelo falante que, ao fazer actuar os seus órgãos articuladores, produz um sinal acústico que é recebido por um ouvinte (Ranchhod 2001: 196). Este sinal é transformado num padrão acústico, para que este seja interpretado pelo sistema de reconhecimento de fala e este o transforme em unidades linguísticas. É ainda importante salientar que existem várias aplicações de reconhecimento que dependem tanto do sinal acústico, como das técnicas de extração do computador: razões que condicionam, por vezes, a funcionalidade destes sistemas.

Por esta razão, “reconhecimento de fala” foi a opção tradutória, para além de ser, obviamente, um termo que já tem uma tradução consagrada e utilizada na área de processamento de línguas naturais, na qual é especialista a professora Elisabete Marques Ranchhod.

19. “**Spoken-language MT** has not yet reached even the stage of real-time testing in non-laboratory settings.” (Hutchins 1999: 13) // “**A TA de língua falada** ainda não alcançou um nível de teste em tempo real, e em cenários não-laboratoriais.”

À semelhança do termo “reconhecimento de fala”, estamos perante um termo que diz respeito ao processamento de língua natural, onde figuram tanto línguas faladas como escritas, sendo que o foco principal incide na língua falada. Embora pareça um processo simples, ao contrário da língua escrita, não se trata de um processo de mera correspondência de texto para texto. “Spoken language is hugely different from written language, apart from the obvious difference of medium (sound vs. text), which involves an amount of pre-linguistic processing to isolate the speech signal from the surrounding background ‘noise’” (Mitkov 2005: 521). Como tinha sido referido anteriormente no exemplo do reconhecimento de fala, este é um processo bastante complexo. Especialmente quando no caso da língua falada – *spoken language* – não estamos perante um sistema, mas de uma tradução automática de língua falada, que é muito mais complexa e problemática.

- Este é outro dos exemplos que ainda não se encontra cunhado no léxico português na área de TA. No entanto, dada a pesquisa encetada, é fácil perceber que estamos novamente a incidir sobre a área de Linguística Computacional e sobre o processamento de línguas naturais. “Língua falada” é um termo sobejamente conhecido na área da Linguística e sabemos, através da bibliografia disponível da Professora Elisabete Ranchhod, que é um termo já cunhado e devidamente catalogado. Existindo, por conseguinte, a diferenciação entre a língua escrita e falada, traduzir este tipo de sistemas de tradução automática como “tradução automática de língua falada” é uma proposta arriscada, porque é pessoal e não se encontra consagrada. Contudo, não sinto que seja uma opção de difícil compreensão, uma vez que foi feita com base em algumas pesquisas bibliográficas em português e inglês que conferem o que foi dito. É ainda importante salientar que existem alguns órgãos que começam, também, a utilizar o termo, como é exemplo o Inesc – Instituto de Engenharia de Sistemas e Computadores, Investigação e Desenvolvimento em Lisboa – que, inclusive, já possui um laboratório de língua falada (Inesc – Instituto de Engenharia de Sistemas e Computadores, Investigação e Desenvolvimento em Lisboa 2013).



20. “Météo is an example of a ‘**sublanguage**’ system, i.e. designed to deal with the particular language of meteorology.” (Hutchins 1999: 6) // “O Météo é um exemplo de um **sistema de ‘sub-linguagem’**, i. e. concebido especialmente para lidar com linguagem específica de meteorologia.”

Especialista em estruturalismo (uma teoria linguística que concebe a língua como uma estrutura, definida pela relação entre os seus elementos), Zellig Harris foi o primeiro a dar uma noção desta teoria em termos linguísticos. Este termo tem uma forte componente matemática, razão pela qual é cunhado com este prefixo, aparentemente estranho. “(...) the mathematical idea of “subsystem” in mind, the “sub-” prefix indicating not exclusion, but inclusion. So a sublanguage is a subset of the “whole” language” (Somers 2003: 283). O prefixo advém, portanto de “subset”, que significa “subconjunto: conjunto cujos elementos também ao conjunto dado; parte de outro conjunto” (Infopedia 2013), que é também em português, uma noção matemática.

Podemos dizer que estamos perante uma sub-linguagem quando um grupo de especialistas comunica entre si, partilhando conhecimento sobre um domínio restrito. Este acto de comunicação será, por conseguinte, constituído por um vocabulário próprio que exhibirá termos técnicos (que derivam da linguagem utilizada no domínio), e palavras coloquiais, do dia-a-dia (que têm uma interpretação narrativa, consoante o domínio específico onde se inserem). Podemos dizer que “(...) When the utterances (including writings) of domain experts show some systematic patterns that distinguish them from the language as a whole, then we say these utterances belong to a sublanguage” (Mitkov 2005: 432).

- De acordo com a bibliografia a que tive acesso, sabia-se que este termo era relativo a um sistema que tinha começado a ser desenvolvido na Universidade de Montreal [Université de Montréal], chegando a mesma a desenvolver aquele que seria o primeiro sistema de “sublanguage” – o *Météo*. Como o nome indica, o *Météo* traduzia a informação e as previsões do tempo de Montreal. Foi possível perceber, também, que ao trabalhar com domínios específicos, o sistema de sub-linguagem demonstrava algumas semelhanças com o sistema de linguagens controladas, uma vez que o léxico e a gramática eram também restritos (em comparação com a linguagem, como um todo). À

semelhança da opção tradutória para “controlled language systems”, também neste termo a noção de “língua” e “linguagem” foi crucial para a aproximação a uma tradução exequível. Nunca poderia figurar algo como “sublíngua”, uma vez que não estávamos perante um subconjunto de um “idioma ou de um sistema abstrato de signos e regras gramaticais” específico. Caso a opção tradutória fosse essa, o leitor poderia concluir que estaria perante uma qualquer língua e que o seu vocabulário era limitado de alguma forma. Dadas as leituras efetuadas, tornou-se evidente que estávamos perante a linguagem no seu todo e no subconjunto de expressões, e não de uma língua em concreto. Linguagem foi, portanto, a escolha final, juntamente com o mesmo prefixo que preconizou o termo. Apesar de ter sido uma opção tradutória ponderada, não foi completamente arriscada, uma vez que se encontraram alguns antecedentes. Não só localizámos a expressão na tese de Ibrahim (Ibrahim 2010: 15), como também em disciplinas de informática lecionadas pelo Instituto Politécnico de Setúbal (IPS – Instituto Politécnico de Setúbal 2013), embora “sub-linguagem” em informática se refira exclusivamente à sub-linguagem de base de dados.

21. “In addition, **many translator workstations** are now offering full automatic translations using MT systems such as Systran, Logos, and Transcend.” (Hutchins 1999: 7) // “Para além disso, muito do **software de trabalho do tradutor** está agora a oferecer traduções automáticas completas, usando sistemas de TA como o Systran, Logos e o Transcend.”

Este é um de muitos termos que não tem qualquer integração no léxico português ou qualquer tradução. A dificuldade imposta recai não só no domínio específico no qual se insere, como na proposta de tradução que tem que ser encarada através do seu significado, e através do possível esclarecimento do leitor.

- “Translator Workstation” como nos diz Hutchins, e como podemos depreender da bibliografia de que dispomos de TA, é um *software* de tradução que foi criado por Martin Kay (conhecido investigador em Linguística Computacional) em 1980. No entanto, foi Alan Melby (outro notável na mesma área) que, à medida que muitas das

tecnologias de linguística computacional eram atualizadas e desenvolvidas, decidiu integrar várias operações e ferramentas neste *software*, com os moldes que lhe são conhecidos nos dias de hoje. É considerado, atualmente, o *software* de tradução mais rentável e mais completo para o tradutor profissional, e para as grandes empresas. Segundo Somers, é composto por ferramentas básicas e por recursos um pouco mais completos, que foram acrescentados e concebidos com o aparecimento de sistemas comerciais. São ferramentas básicas as operações de processamento de palavras, dicionário de sinónimos, um corretor ortográfico, número de palavras, e ainda uma base de dados terminológica. Mais tarde, com o aparecimento da *internet*, foram acrescentadas funcionalidades para serem acedidas virtualmente, como dicionários e terminologia *online*. No entanto, uma das operações de destaque deste *software*, é o facto de permitir o alinhamento dos dois textos de trabalho através de outro *software* de tradução: “translation memory” [a memória de tradução]. Oferece, assim, a possibilidade ao tradutor humano de uma tradução interativa, na qual o sistema interrompe algumas funções para perguntar ao tradutor qual a operação que prefere (Somers 2003: 23).

“Translation Workstation” é um termo que pode induzir o leitor em erro, dado que muitas das funcionalidades que este *software* possui são frequentemente descritas por Hutchins como “facilities”. Pode ser um erro muito grave, caso não haja rigor terminológico aquando da leitura e tradução. Existem dois motivos pelos quais surge uma dificuldade acrescida de compreensão do termo: não só o seu nome é, através de uma tradução literal, bastante dúbio, como as operações em que consistem caracterizam-se através de um substantivo que tem mais que um significado. Começemos então pelo nome deste termo: “Workstation” ou “Workbench” foram dois dos nomes dados a este *software*. Workstation é um nome que designa um “posto de trabalho; área de trabalho; terminal de computador” (Porto Editora 2009: 1327) e Workbench, algo que designa uma “banca de trabalho; bancada de trabalho” (Porto Editora 2009: 1327). Ambos referem-se, portanto, especificamente a um local de trabalho. O leitor que tenha um conhecimento proficiente da língua inglesa deduz, por conseguinte, que se trata de um local ou infraestrutura na qual o tradutor trabalha. Esta assunção é coadjuvada pela qualificação de muitos dos seus serviços como “facilities”, que significa “facilidade, habilidade; função, recurso; instalações; serviços” (Porto Editora 2009: 311). Por esta razão, figurava inicialmente a proposta de tradução “local

de trabalho do tradutor”, caindo-se no erro de assumir que se tratava de instalações de trabalho do tradutor que possuem as ferramentas descritas, e não um *software* de trabalho.

Contudo, após a leitura da descrição de Hutchins e de outros autores da nossa bibliografia, pudemos finalmente compreender que não se trata especificamente de um local de trabalho, mas de um programa de *software* que é composto por um conjunto de ferramentas. O nome pode, quando analisado, resultar da autonomia que é conferida ao tradutor através de todas estas ferramentas, que permitem que o *software* seja encarado como um posto de trabalho completo. No entanto, pelas razões descritas acima, a tradução literal não seria ponderada. Se o termo em inglês induz em erro qualquer leitor, neste caso em português uma opção tradutória à letra é, definitivamente uma má escolha. Uma vez que a intenção é fazer com que o leitor assimile a maior quantidade de informação possível com o menor esforço (Byrne 2006: 7) então, a proposta de tradução para um leitor de língua portuguesa tem que ser algo que contenha a simbologia mas, essencialmente, o conteúdo que lhe deu origem. Por esta razão, e dado o rigor exigido, a opção proposta é “*software* de trabalho do tradutor”. Caso figurasse algo como “*software* de várias ferramentas de trabalho do tradutor” era algo que, embora bastante elucidativo, era demasiado extenso. Dado que a palavra *software* é já um empréstimo consagrado pelo uso, e compreende já um conjunto de instruções ou programas que possibilitam a leitura de informação e o funcionamento do computador, a opção sugerida indica um programa que contém ferramentas indispensáveis ao trabalho do tradutor.

22. “The translator can thus search for a phrase or even full sentence in one language in the **translation memory** and have displayed corresponding phrases in the other language.” (Hutchins 1999: 7) // “O tradutor pode assim pesquisar uma expressão, ou até mesmo uma frase inteira numa língua na **memória de tradução**, e ter em exibição as expressões correspondentes numa língua e noutra.”

A memória de tradução era, inicialmente, uma das ferramentas integradas no *software* de trabalho do tradutor. Embora só se tivesse tornado comercialmente

disponível no mercado a partir dos anos 90 do século XX, esta ferramenta fora inicialmente proposta nos anos 70 do mesmo século. Não se sabe ao certo a quem atribuir a autoria da sua proposta no mercado da TA, se a Martin Kay, Alan Melby ou David Arthern. De acordo com vários registos, tanto a memória de tradução como o *software* de trabalho do tradutor foram ideias originais que surgiram da publicação de Kay: *The Proper Place of Men and Machines in Language Translation*. No entanto existem registos de uma sugestão semelhante dada pelo tradutor Peter Arthern<sup>16</sup>. Esses registos dizem que esta sugestão precede a de Kay, na qual Arthern descreve uma funcionalidade de acesso *online* a documentos já traduzidos, que corresponde às funcionalidades de uma memória de tradução. Arthern afirma “It must be possible to produce a programme [sic] which would enable the word processor to ‘remember’ whether any part of a new text typed into it had already been translated, and to fetch this part, together with the translation which had already been translated (...)” (Arthern *Apud* Somers 2003: 33). Porém, também Alan Melby afirma ser um dos especialistas que, com o seu grupo de estudo na Brigham University [Universidade de Brigham] pensou numa solução semelhante de auxílio ao tradutor. Independentemente de quem surgiu com a ideia, todos concordavam num ponto essencial: ao programa ou sistema era imprescindível uma base de dados de exemplos de tradução. Existem três formas de criar uma base de dados: construindo-a à medida que se desenvolve trabalho, importando de outro local ou criando-a através de um texto paralelo (Somers 2003: 35). Qualquer que seja a opção escolhida, é necessário que a memória de tradução desempenhe uma das suas funções mais importantes: a capacidade de corresponder a frase a ser traduzida com a base de dados que possui.

- Este não é um termo que suscite grandes dificuldades de compreensão por parte do leitor, ou uma grande dificuldade tradutória do termo. Como o nome indica, “translation memory” é um *software* de tradução cuja funcionalidade principal consiste em consultar uma base de dados de traduções previamente feitas, que são memorizadas no *software*. O intuito é facilitar a consulta do tradutor que, ao procurar algo semelhante ao que pretende traduzir na base de dados, pode utilizar esse exemplo como um modelo ou

---

<sup>16</sup> Existe muito pouca informação sobre Peter Arthern, o que impossibilitou que fosse acrescentado algo a seu respeito. A única bibliografia disponível refere-se exatamente à sua citação que, segundo se diz, precede a de Martin Kay sobre a ideia original sobre a Memória de Tradução.

sugestão de como determinada palavra deve ser traduzida (Somers 2003: 31). Um dos objetivos do programa é memorizar e armazenar a maior quantidade possível de frases na sua base de dados, de forma a melhorar a eficácia do sistema de correspondência de frases que o *software* possui. É importante ressaltar que este termo tem duas utilizações: tanto pode ser utilizado para descrever o *software* – referido como “sistema de memória de tradução” – como para denominar a própria base de dados, i.e., a memória que possui de traduções prévias. Embora grande parte da terminologia de TA exija que as propostas de tradução pensadas sejam uma alternativa não-literal de muitos dos termos, este é um dos poucos exemplos nas quais uma tradução à letra não constitui, de todo, qualquer entrave. Neste caso, o nome do termo espelha exatamente as funcionalidades que possui, pelo que a tradução literal não se torna errónea ou obtusa para o leitor. E, assim, este consegue perceber o que se pretende deste *software*. É ainda importante notar que é dos poucos termos do nosso glossário que, tanto na área de TA, como na área de tradução em geral já se tornou consagrado pelo uso. A título de exemplo temos uma publicação do vocabulário em português existente nas memórias de tradução da Comissão Europeia (Ec.europa 2013), os recursos de tradumática do Instituto da Cooperação e da Língua, Camões Portugal (Cvc.instituto-camoes 2013)<sup>17</sup> ou a AP Portugal, conhecida empresa de tradução (Apportugal 20013)<sup>18</sup>.

23. “Systran Pro and Systran Classic, for example, are **Windows-based versions** of the successful system developed since the 1960s for clients worldwide in a large range of languages;” (Hutchins 1999: 9-10)// “O Systran Pro e o Systran Classic, por exemplo, são versões baseadas **no sistema operativo Windows [Windows-based]**, do sistema de sucesso desenvolvido desde 1960 para clientes em todo o mundo numa grande variedade de línguas.”

Este é uma expressão que encontrei apenas na publicação de Hutchins. A língua inglesa tem, ao contrário da língua portuguesa, estruturas que permitem construções nas quais um simples nome adjetiva e qualifica um termo, cunhando-o. Em inglês, o acréscimo de nomes, adjetivos, ou até prefixos e sufixos é suficiente para qualificar uma

---

<sup>17</sup> In <http://www.instituto-camoes.pt/> (2013)

<sup>18</sup> In <http://www.apportugal.com/> (2013)

palavra, tornando-a rapidamente num termo específico, cuja compreensão é geralmente bastante simples e fácil. Como é possível depreender, estamos perante versões do sistema Systran que se baseiam no sistema operativo Windows e nas suas funcionalidades. Como é sabido, o Windows é um sistema operativo mundialmente conhecido, razão pela qual dispensa apresentações. É um sistema que toda a gente conhece, embora não saiba descrever de forma correta. Isto porque qualquer PC – Computador pessoal, é equipado com este sistema operativo (à exceção, claro, de um Macbook© que possui o sistema Mac OS). O Windows tem inúmeras versões que, ao longo dos anos, são atualizadas para que a empresa consiga manter a liderança do mercado. Apesar das constantes mudanças, existem programas que o Windows sempre manterá, como o Word, Excel, PowerPoint, entre outros, que qualquer utilizador conhece.

- Por ser um termo provavelmente cunhado pelo próprio Hutchins, foi difícil encontrar qualquer tipo de informação que suportasse qualquer opção tradutória. Por esta razão, e partindo do princípio que estamos perante um termo que o próprio designou, decidimos analisar o contexto em que se inseria para proceder, assim, à sua tradução. Como foi referido, estamos perante um sistema operativo que qualquer utilizador conhece sobejamente. Por se tratar de um nome vulgarmente conhecido, o nome era insubstituível na opção tradutória. Se o sistema Systran e as suas versões foram baseados num sistema como o Windows, existiam apenas duas opções tradutórias: apenas “versões baseadas no Windows” ou “versões baseadas no sistema operativo Windows”. Apesar de ambas as versões serem claras, a opção final escolhida foi a última. Isto porque “versões baseadas em Windows” era uma escolha demasiado coloquial e poderia implicar que seria necessário indicar qual a versão do Windows, na qual o Systran se baseava. Por esta razão, optei por incluir “sistema operativo”, elucidando o leitor (mesmo que essa preocupação não seja necessária) de que se tratava de um sistema operativo, e que o Systran baseia, portanto, as suas funcionalidades nas do sistema operativo que mencionam – independentemente da versão utilizada.

**c) Respeito pelas convenções do texto técnico**

Tentei elaborar a proposta de tradução, não esquecendo o tipo de texto que a publicação de Hutchins era. Dado que estamos perante um texto técnico, a proposta de tradução tinha que ser tão objetiva e clara para o leitor quanto o texto de partida. Por motivo optei por não entrar em grandes desvios em relação à forma original.

Tentei, por isso, ser fiel à forma como o autor apresenta o conteúdo, e construí a informação. Procurei o máximo de objetividade (e a máxima correspondência terminológica entre as duas línguas) requerida pela natureza do texto em causa.



## Considerações Finais

A exequibilidade deste trabalho passa não só pela seleção criteriosa dos textos que apresenta, mas também pela metodologia adotada. A utilização desta metodologia permitiu delinear estratégias de tradução para enfrentar possíveis problemas de tradução, e veio também conferir outra perspectiva sobre os textos desta área.

O modelo de análise textual de Christiane Nord permitiu pôr em evidência, não só as características principais do texto, mas também as que normalmente são consideradas secundárias e que no modelo da autora têm a mesma importância. Isto veio conferir outra perspectiva crítica sobre o seu conteúdo, e permitiu ainda observar a importância dos elementos extratextuais para a análise do texto de partida e consequente tradução. A aplicação do modelo de Nord possibilitou ainda que fossem definidas opções de tradução para tentar solucionar os problemas que surgissem. A análise aos textos selecionados obedeceu, por conseguinte, à metodologia proposta e as opções que figuram nas propostas de tradução são, portanto, pessoais. Embora possa ser possível encontrar lacunas e/ou outras alternativas de análise e tradução, as apresentadas respeitaram a análise textual presente no modelo de Nord.

A aplicação deste modelo aos textos escolhidos pode, a meu ver, ajudar o leitor a compreender o quão complexa a área da TA pode ser, e de como fatores como a cultura de partida e chegada, o espaço, o léxico, entre outros, podem ser elementos decisivos para uma proposta de tradução viável e para a futura compreensão do leitor. A utilização deste modelo ajudou também a compreender a complexidade e a especificidade do tipo de linguagem utilizada na área da TA, que foi também uma das motivações para este trabalho.

Os problemas identificados e as propostas de tradução apresentadas em anexo podem interessar a alunos de Tradução, ou de áreas semelhantes, que pretendam ter acesso a uma informação mais detalhada sobre o tema e queiram compreender melhor tudo o que diga respeito à TA. Como já foi referido, existem várias obras disponíveis que o aluno pode pesquisar, tanto em inglês como em português. Em língua portuguesa existem várias obras, como as que foram mencionadas no breve relatório de investigação (ponto 1.2. da Introdução). Embora seja mais fácil ter acesso a estes

conteúdos em inglês, torna-se uma tarefa mais difícil e um pouco mais trabalhosa, para um leitor português, decifrar a informação apresentada.

É por este motivo que este trabalho visa ser uma contribuição útil no contexto académico, que através da metodologia adotada e das propostas de tradução que apresenta, pode servir como uma ferramenta de pesquisa que pretende facilitar o acesso à informação e à compreensão desta complexa temática.

## Referências Bibliográficas

### 5.1. Bibliografia Geral

S/A (2013) *Dicionário da Língua Portuguesa - com Acordo Ortográfico*, Porto, Porto Editora, edição de 2013.

S/A (2013) *Dicionário de Inglês-Português*, Porto, Porto Editora, edição de 2013.

S/A (2009) *Dicionário de Inglês-Português*, Porto, Porto Editora, 5ª edição.

Almeida, José João; Simões, Alberto (2007) "XML::TMX - Processamento de Memórias de Tradução de Grandes Dimensões", in José Carlos Ramalho; João Correia Lopes; Luís Carriço (ed.), *XML: Aplicações e Tecnologias Associadas (XATA2007)* (FCUL, Lisboa, 15-16 de Fevereiro), Universidade do Minho, pp. 83-93.

Baker, Mona (2006) *In Other Words: A Coursebook On Translation*, New York, Routledge, 2ª edição.

Baker, Mona (2006) *Routledge Encyclopedia of Translation Studies*, New York, Routledge, 1ª edição.

Barrento, João (2002) *O Poço de Babel : Para uma Poética da Tradução Literária*, Lisboa, Relógio D'Água, 2ª edição.

Byrne, Jody (2006) *Technical Translation, Usability Strategies for Translating Technical Documentation*, Dordrecht, The Netherlands, edição de 2006.

Correia, Margarita (2005) “A Terminologia em Portugal e países de língua portuguesa em África”, in revista *Terminómetro*, número especial: A terminologia em Portugal e países de língua, pp. 15-20. Acedido a 20 de Setembro de 2013, em: <http://www.iltec.pt/pdf/wpapers/2005-mcorreia-terminometro1.pdf>.

Curado, J. M. (2000) *O mito da tradução automática*, Braga, Universidade do Minho/Centro de Estudos Humanísticos, Coleção Hespérides.

Fontanet, Mathilde (2005) “La Traduction Technique: le texte sous l’empire de l’extratextuel”, in *Práctes des 7es Journées Scientifiques AUF-LTT, Mots, Termes et Contextes*, Genève, Acedido a 23 de Setembro de 2013, em: <https://archive-ouverte.unige.ch/unige:14834>.

Gambier, Yves e Doorslaer, Luc Van (2010) *Handbook of Translation Studies*, Amesterdão, John Benjamins Publishing Company, edição de 2010.

Gomes, Ana Lucrecia Madeira (2010) *Tradução Automática e Linguagens Controladas: Contributos para um Português Controlado*. Dissertação de Mestrado, Faculdade de Letras da Universidade de Lisboa.

Krieger, Maria; Finatto, Maria José (2004) *Introdução à Terminologia, teoria & prática*, São Paulo, editora Contexto.

Mailliot, Jean (1975) *A Tradução Científica e Técnica*, Brasília, Brasil, editor McGraw-Hill do Brasil Ltda.

Mikel L. Forcada (2010) "Machine Translation Today", in Yves Gambier, Luc Van Doorslaer, ed., *Handbook of Translation Studies*, vol. 1, pp. 215-223.

Mitkov, Ruslan (2005) *The Oxford Handbook of Computational Linguistics*, USA, Oxford University Press.

Munday, Jeremy (2009) *The Routledge Companion to Translation Studies*, Oxon, Routledge,

Náheda, Ibrahimo (2010) *Para uma tradução automática baseada em conhecimento: especificação da modificação e predicação adjectiva*. Dissertação de Mestrado, Faculdade de Letras da Universidade de Lisboa.

Neves, Jorge (2009) *Utilizar o Computador, Depressa e Bem*, Lisboa, FCA, 7ª edição.

Nirenburg, Sergei (2003) *Readings in Machine Translation*, Cambridge, Massachusetts, London, England, Bradford Book, 1ª edição.

Nord, Christiane (2005) *Text Analysis in Translation. Theory, Methodology, and Didactic Application of a Model for Translation-Oriented Text Analysis*, Netherlands, Rodopi Editions, 2ª edição.

Ranchhod, Elisabete Marques (2002) *Advances in Natural Language Processing, Third International Conference, PorTal 2002 Faro, Portugal, June 2002, Proceedings*, Germany, editora Springer.

Ranchhod, Elisabete Marques (2001) *Tratamento das Línguas por Computador, Uma introdução à Linguística Computacional e suas aplicações*, Lisboa, Editorial Caminho, 2ª edição.

Santos, Diana (2007) *Avaliação conjunta: um novo paradigma no processamento computacional da língua portuguesa*, Lisboa, IST Press.

Santos, Diana (2006) “A tradução na sociedade do conhecimento ou Tradução: uma tecnologia humana de ponta ou Ciência e tradução”, in *IX Seminário de Tradução Técnica e Científica*, Lisboa. Acedido a 15 de Setembro de 2013, em: <http://www.linguateca.pt/Diana/download/SantosSeminTradCiencia2006actas.pdf>.

Seruya, Teresa (2013) “Anthologies and translation”, in Seruya, Teresa; D’Hulst, Lieven & Moniz, Maria Lin. Teresa Seruya et al. (Eds.) *Translation anthologies and collections: an overview and some prospects (19th and 20th Centuries)*, Amsterdam/Philadelphia: John Benjamins.1-16.

Simões, Alberto Manuel Brandão (2008) *Extracção de Recursos de Tradução com base em Dicionários Probabilísticos de Tradução*, Dissertação de Doutoramento, da Faculdade de Engenharia da Universidade no Minho. Acedido a 15 de Setembro de 2013, em:  
<http://www.linguateca.pt/documentos/SimoesPHD.pdf>.

Somers, Harold (2003) *Computers and Translation, A translator’s guide*, Amsterdam/Philadelphia, John Benjamins Publishing Company.

Vilela, Mário (1999) *Gramática da Língua Portuguesa*, Lisboa, Edições Almedina, 2ª edição.

## 5.2. Recursos online

BAD, Associação Portuguesa de Bibliotecários, Arquivistas e Documentalistas - <http://www.bad.pt/publicacoes/index.php/congressosbad/article/view/334>, acedido a 30 de Novembro de 2013

Camões, Instituto da Cooperação e da Língua - <http://www.instituto-camoes.pt/>, acedido a 6 de Novembro de 2013

CLUL, Centro de Linguística da Universidade de Lisboa - <http://www.clul.ul.pt/pt>, acedido a 13 de Setembro de 2013

FLUL, Faculdade de Letras da Universidade de Lisboa - <http://www.fl.ul.pt/dlgr-docentes/913-dlgr-programas-traducao-assistida-por-computador>, acedido a 10 de Novembro de 2013

*Iltec, Instituto de Linguística Teórica e Computacional* - <http://www.portaldalinguaportuguesa.org/?action=terminology>, acessado a 20 de Outubro de 2013

*INESC – Instituto de Engenharias de Sistemas e Computadores* - [https://www.l2f.inesc-id.pt/wiki/index.php/Pesquisas\\_em\\_PLN\\_no\\_N%C3%BAcleo\\_de\\_Lingu%C3%ADstica\\_Computacional\\_\(NILC\)](https://www.l2f.inesc-id.pt/wiki/index.php/Pesquisas_em_PLN_no_N%C3%BAcleo_de_Lingu%C3%ADstica_Computacional_(NILC)), acessado a 20 de Novembro de 2013; [https://www.l2f.inesc-id.pt/wiki/index.php/Main\\_Page](https://www.l2f.inesc-id.pt/wiki/index.php/Main_Page), acessado a 21 de Novembro de 2013

*Infoescola, Trabalhos Acadêmicos e Pesquisas Escolares, cursos online* - <http://www.infoescola.com/>, acessado a 2 de Outubro de 2013

*Infopedia, Enciclopédia e Dicionários Porto Editora* - <http://www.infopedia.pt/>, acessado a 16 de Outubro de 2013

*IPS, Instituto Politécnico de Setúbal* - [http://www.si.ips.pt/ests\\_si/Disciplinas\\_geral.FormView?P\\_ANO\\_LECTIVO=2013/2014&P\\_CAD\\_CODIGO=INF21115&P\\_PERIODO=15](http://www.si.ips.pt/ests_si/Disciplinas_geral.FormView?P_ANO_LECTIVO=2013/2014&P_CAD_CODIGO=INF21115&P_PERIODO=15), acessado a 17 de Outubro de 2013

*IST, Instituto Superior Técnico* - <https://fenix.ist.utl.pt/disciplinas/ln-2/2012-2013/1-semester/planeamento>, acessado a 13 de Setembro de 2013

*Label, Laboratório de Engenharia da Linguagem* - <http://label.ist.utl.pt/pt/apresentacao.php>, acessado a 28 de Setembro de 2013

*Léxico Técnico Português: Ambiente, Banca, Comércio, Construção, Energia, Seguros, Turismo, Telecomunicações, Direito Comercial Internacional e Economia e Gestão de Empresas*, acessado a 13 de Setembro de 2013, em:  
<http://instituto-camoes.pt/lextec/>

*Linguateca, centro de recursos para o processamento computacional da língua portuguesa* - <http://www.linguateca.pt/>, acessado a 20 de Setembro de 2013

*Ministério das Finanças, Direção Geral do Tesouro e Finanças* - [http://www.dgtf.pt/ResourcesUser/SEE/Documentos/see\\_ana/ana\\_12\\_06\\_2012\\_documentos\\_confidenciais.pdf](http://www.dgtf.pt/ResourcesUser/SEE/Documentos/see_ana/ana_12_06_2012_documentos_confidenciais.pdf), acedido a 13 de Setembro de 2013

*Priberam, Dicionário Priberam da Língua Portuguesa* - <http://www.priberam.pt/dlpo/>, acedido a 20 de Setembro de 2013

*Scribd, The World's Largest Digital Library* - <http://pt.scribd.com/doc/135333247/jurisprudencias-desistencia-desclassificacao>, acedido a 13 de Setembro de 2013

*Visibilidade, Glossário de Informática e tradução de termos informáticos* - <http://visibilidade.net/tutorial/glossario-informatica.html>, acedido a 20 de Setembro de 2013

*York University* - [http://www.psych.yorku.ca/gigi/documents/Thorndike\\_1943.pdf](http://www.psych.yorku.ca/gigi/documents/Thorndike_1943.pdf), acedido a 25 de Novembro de 2013



## **Anexo 1: *Translation*, de Warren Weaver**